



**Dottorato di ricerca in Scienze della Terra – Ciclo XVIII**

**Sede: Università degli Studi di Napoli Federico II**

**Antonio Esposito**

**SVILUPPO DI PROCEDURE AUTOMATIZZATE IN AMBIENTE  
GIS PER L'ANALISI DI FORME TERRAZZATE E SUPERFICI  
GEOMORFOLOGICHE RELITTE**

**Il Tutore**

**Prof. Aldo Cinque**

**Il Coordinatore**

**Prof. Giuliano Ciampo**

**Gennaio 2007**

## **Sommario**

<b>Ringraziamenti</b>	<b>i</b>
<b>Premessa</b>	<b>ii</b>
<b>Struttura della tesi</b>	<b>iii</b>
 <b>Parte A – Introduzione</b>	
<b>A.I – Obiettivi della tesi</b>	<b>2</b>
A.II – Le superfici di maggior interesse	4
1.1 Terrazzi fluviali	4
1.2 Glacis d'erosione in roccia tenera	11
1.3 Altre superfici relitte	13
 <b>Parte B – Dati di letteratura</b>	
<b>B.I – Concetti di base</b>	<b>17</b>
2.1 Il Gis	17
2.2 Approccio metodologico	18
2.3 La geomorfologia quantitativa	19
2.4 La Geomorfometria	21
2.5 Metodi di analisi tradizionale (quantitativa e qualitativa)	24
2.5.1 Analisi di foto aeree	24
2.5.2 Carte topografiche	25
2.5.3 Analisi del profilo topografico	26
<b>B.II – Modelli digitali della superficie topografica</b>	<b>31</b>
3.1 Introduzione	31
3.2 Metodi di rappresentazione ed archiviazione dei DTM	31
3.3 Cenni sui metodi di acquisizione	39
3.4 Metodi di interpolazione per la generazione di DTM	44
3.4.1 Metodi globali	46
3.4.2 Metodi locali	53
3.4.3 Metodi geostatistici: il kriging	55
3.5 Parametri morfometrici estraibili dall'analisi differenziale	58
3.6 Parametri determinati dal flusso idrico	64
3.7 Qualità dei DTM	68
<b>B.III – Operazioni sui raster</b>	<b>72</b>
4.1 Il modello raster	72
4.2 I filtri di convoluzione	76
4.2.1 Filtri gradiente	79
4.2.2 Filtri di smoothing	82
4.2.3 Filtro laplaciano	84
4.3 Neighborhood Analysis	85
4.4 Lo shaded relief	86
4.5 Le operazioni locali e la map-algebra	88
4.6 Le operazioni zonali	90

B.IV – Classificazione del pattern	92
5.1 Introduzione	92
5.2 Definizione di pattern	94
5.3 Tecniche di Classificazione supervisionata	95
5.4 Tecniche di Classificazione non supervisionata	97

## **Parte C – Elaborazioni sulle aree test**

C.I – Introduzione	106
C.II – Le aree test	109
6.1 L'alta valle del fiume Volturno	109
6.2 La valle del fiume Trigno	116
C.III – Elaborazioni sui raster	120
7.1 Precedenti conoscenze	120
7.2 Il metodo sviluppato per la classificazione non supervisionata	124
7.3 Cartografazione supervisionata	157
7.4 Correlazione ed ordinamento delle superfici	165
7.5 Algoritmo sviluppato	169
7.6 Correlazione ed ordinamento dei lembi di terrazzi fluviali	175
7.7 Analisi ed ordinamento dei glaci	178
C.IV - Conclusioni e sviluppi futuri	185
<b>Bibliografia</b>	189
<b>Appendice</b>	197

## ***Ringraziamenti***

*I miei più sentiti ringraziamenti vanno innanzitutto ai miei tutor, professor Aldo Cinque e Dottor Pietro Aucelli, per avermi seguito pazientemente in questi anni, durante le difficili fasi della ricerca e la stesura dell'elaborato di tesi. Grazie alla loro esperienza ho potuto ampliare le mie conoscenze in campo geomorfologico. Vorrei ringraziare diverse altre persone. In particolare un cordiale ringraziamento al Dottor Antonio P. Leone, per la sua preziosa revisione critica del testo e l'Ingegnere Carmine Maffei, il quale mi ha fornito utili consigli su diversi aspetti matematici dell'elaborato come pure il professor Ciro Marmolino per i suoi consigli in materia. La professoressa Carmen Roskopf, per aver contribuito alla revisione del testo ed ampliato le mie conoscenze sulle aree di studio. Credo che tra tutte le persone che ho avuto il piacere di conoscere, il professor Ludovico Brancaccio sia stata una delle poche che oltre ad essere infinitamente paziente e di una cordialità ineguagliabile, è stata in grado di suscitare sempre vivo interesse per lo studio e l'approfondimento. La sua grande capacità di trasmettere le conoscenze agli allievi mi ha alla fine portato verso lo studio delle forme della terra. A lui i miei più vivi ringraziamenti anche per la revisione critica del testo.*

*Ringrazio tutta la mia famiglia per avermi sempre aiutato nei momenti più difficili e a cui doverosamente dedico il mio lavoro.*



## ***Premessa***

La presente ricerca di tesi è incentrata sulla individuazione e sperimentazione di alcune metodologie basate sull'uso del GIS a supporto delle analisi geomorfologica di tipo tradizionale. In particolare ci si è posti l'obiettivo di individuare tecniche di analisi finalizzate alla cartografazione automatizzata di forme considerate di fondamentale importanza per lo studio dell'evoluzione di medio e lungo termine del paesaggio. Nei vari capitoli verranno descritte sia le metodologie utilizzate per le elaborazioni originali, che metodologie note in letteratura utilizzate in vari lavori attinenti l'argomento della tesi. Alcune di queste tecniche di analisi risultano già da tempo implementate a livello software in diversi pacchetti GIS sia commerciali che di libera distribuzione. Per quanto riguarda le metodologie sperimentate è stato invece necessario modificare algoritmi già esistenti (che implementano tecniche di elaborazione descritte in letteratura) ma non pienamente integrati nel software GIS utilizzato (ad esempio l'analisi multivariata descritta nel paragrafo 5), o crearne di completamente nuovi (algoritmi ad hoc), sviluppati appositamente per risolvere problemi più specifici della presente ricerca. Allo scopo di individuare quale supporto era già possibile ottenere dai pacchetti GIS esistenti e più in generale dai metodi di analisi automatica già implementati, sono state quindi esaminate le tecniche utilizzate da diversi autori, specialmente quelle derivanti dall'area di studio dell'analisi ed il riconoscimento automatico delle forme\*. Ci si è innanzitutto resi conto del fatto che numerosi metodi di analisi automatica della superficie terrestre sono nati come necessità derivanti da campi scientifici diversi da quelli delle Scienze della Terra (ingegneria, analisi delle immagini, statistica ecc..) ma che poiché negli anni sono stati oggetto sempre maggiore dell'attenzione di chi si

---

\* benché la traduzione in “riconoscimento delle forme” del termine inglese “pattern recognition” sia quella spesso citata nella letteratura italiana, il termine forme può risultare fuorviante nel caso del presente lavoro. Si utilizzerà quindi in avanti il termine forma per indicare concetti geometrici (es. forma del versante) ed il termine inglese *pattern* per l'analisi multivariata. Si rimanda per la definizione di *pattern* al capitolo 5.

occupa della geomorfologia dal punto di vista quantitativo, sono stati via via inglobati in ambiente GIS. Dopo aver individuato le metodologie di analisi più utili alla caratterizzazione del rilievo, si è valutata la loro efficacia ai fini della cartografazione automatica. Si sono quindi scelte alcune aree test per sperimentare approcci noti e per sviluppare le esperienze acquisite in tecniche originali (descritte nella parte C.III) ed implementarle a livello software.

### ***Struttura della tesi***

Il testo è suddiviso in quattro parti **A**, **B**, e **C** ed un'appendice.

La parte **A** comprende l'introduzione con gli obiettivi della tesi.

La parte **B** è incentrata sulla descrizione dei concetti di base: vengono descritte le principali tecniche di analisi automatica della superficie terrestre attraverso l'utilizzo dei modelli digitali del terreno da dati di letteratura. Molte delle tecniche di analisi e dei concetti illustrati compaiono in pubblicazioni che riguardano sia applicazioni specifiche del campo delle Scienze della Terra che di campi completamente diversi quali il riconoscimento del *pattern* o l'analisi delle immagini digitali.

Nella parte **C** si descrivono le aree test, gli algoritmi proposti e le elaborazioni effettuate sulle aree test tramite i medesimi.

Una appendice in cui sono inclusi i principali segmenti di codice sorgente degli algoritmi sperimentati per la cartografazione automatica.

*Parte A*  
*Introduzione*

## ***A.I Obiettivi della tesi***

Nonostante le tecniche di cartografazione automatica proposte risultino alla fine utilizzabili per la classificazione del rilievo in contesti morfologici generici, ci si è soffermati in particolare sulla possibilità di evidenziare alcune tipologie di forme, caratteristiche delle aree test impiegate per sperimentare gli algoritmi, originatesi in contesti morfoclimatici lontani nel tempo. Alcune forme si rivelano *marker* utili all'analisi delle fasi morfoevolutive del paesaggio e quindi di particolare utilità per il geomorfologo interessato allo studio dell'evoluzione di medio e lungo termine. Esempi di *marker* di questo tipo sono certamente i terrazzi fluviali, che testimoniano l'evoluzione del fondovalle e la cui analisi consente di ottenere una ricostruzione sufficientemente fedele dell'evoluzione e della morfologia dell'antico sistema fluviale nell'arco di tempo in cui ricade la generazione di tali forme. Altre superfici di interesse sono i glacis d'erosione in roccia tenera e le superfici di spianamento sommitali.

Lo studio dei rapporti tra superfici relitte e fasi morfoevolutive comporta in molti casi la necessità di superare tutta una serie di problematiche dovute alla complessità intrinseca di tali relazioni.

Lo scopo iniziale della tesi, inizialmente confinato al riconoscimento di tali tipologie di forme, si è esteso all'analisi e riconoscimento meno mirato e più generalizzato, ovvero alla suddivisione dei DTM in forme del paesaggio, intese come aree omogenee per caratteristiche morfometriche. In particolare si è cercato di sviluppare una metodologia che consentisse di realizzare in automatico una cartografazione il più possibile paragonabile, dal punto di vista del significato della suddivisione in classi, a quella realizzata tramite metodi manuali. In tale contesto si inserisce il problema più complesso della cartografazione dei terrazzi fluviali, caratterizzati da più ordini di generazioni. Risulta utile ai fini della comprensione del perché sono stati sviluppati algoritmi differenti, riassumere le principali problematiche affrontate:

- Riconoscimento e cartografazione automatica di forme del paesaggio secondo criteri geomorfologici: in particolare le forme descritte in precedenza.
- Correlazione automatica degli ordini di terrazzi fluviali e lembi di superfici relitte e conseguente suddivisione in ordini di generazioni.
- Analisi dei risultati ottenuti e loro validazione.

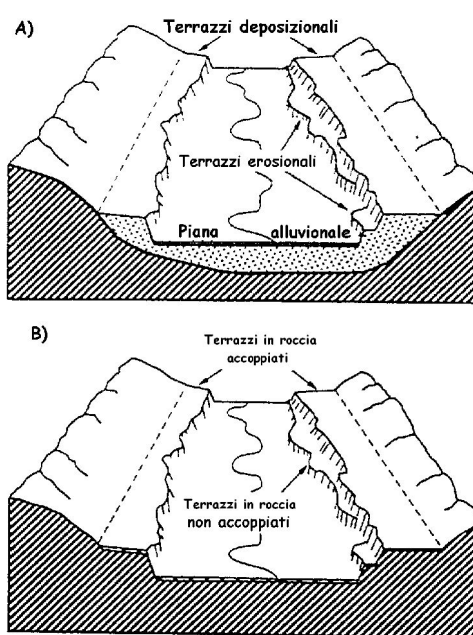
## ***A.II Le superfici di maggior interesse***

### **1.1 Terrazzi fluviali**

Un terrazzo fluviale è una forma di origine fluvio-lacustre testimone di un antico livello del fondovalle ed è il risultato dell'approfondimento del reticolo idrografico ed erosione laterale delle sue sponde (Castiglioni 1986, Panizza 1992). I terrazzi fluviali possono essere sia di tipo deposizionale che di tipo erosionale (fig 1.1). I terrazzi deposizionali sono costituiti dai depositi alluvionali intagliati ed erosi lateralmente. I terrazzi erosionali possono formarsi sia per erosione dei depositi fluviali che essere intagliati in roccia (*strath terraces* sensu Merritts et alii 1994). I terrazzi su roccia sono sempre di origine erosionale e si formano dopo che l'erosione laterale da parte del corso d'acqua, lungo i fianchi vallivi in roccia, ha portato allo svasamento di una valle a V. Più cicli di approfondimento dell'alveo e successiva erosione laterale delle sponde generano livelli terrazzati interamente costituiti da roccia in posto. Durante la fase di erosione laterale agisce lo scalzamento al piede delle scarpate fluviali del fondovalle attivo; questo sia nel caso dei terrazzi in roccia che nel caso dei ripiani formati da depositi alluvionali. Nel caso dei terrazzi formati da depositi alluvionali, si dicono terrazzi erosionali i ripiani intagliati su depositi coevi (fig 1.2a). Successive fasi di riempimento possono determinare la formazione di terrazzi deposizionali più recenti, a contatto con i depositi più antichi (fig 1.2b). Sia per i terrazzi in roccia che per i ripiani intagliati su depositi, è possibile che l'erosione agisca in maniera asimmetrica lungo le due sponde, cancellando un ripiano di terrazzo su di una e conservandolo su quella opposta, producendo terrazzi disaccoppiati (fig 1.1).

Come è noto, uno dei passaggi obbligati (e cruciali) attraverso cui passa una analisi geomorfologia che riguardi dei relitti di antichi ripiani è quello che vede – a valle della individuazione dei singoli lembi relitti - i tentativi di correlarli, tutti

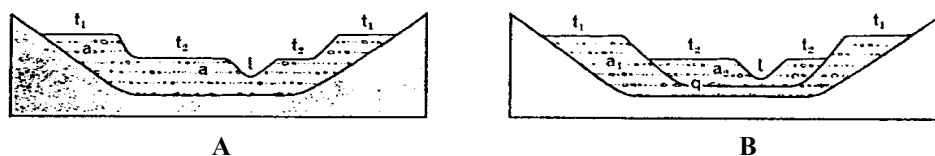
insieme o per gruppi (ordini). Ritenere due lembi di ripiani tra loro correlabili significa ritenerli coetanei; almeno nel senso di appartenenti alla stessa fase morfoevolutiva, che può essere stata più o meno lunga ed aver visto il modellamento dei ripiani più o meno diacrono attraverso l'area considerata. Correlare tra loro dei lembi relitti di ripiani significa anche, e soprattutto, riconoscerne la “accordanza altimetrica” ossia il loro allinearsi l'uno sulla proiezione dell'altro. La accordanza dei terrazzi può essere studiata tramite dei profili longitudinali (distanze-altezze) ed avendo i lembi rappresentati da segmenti isolati si riterranno accordanti quei lembi che, estrapolati l'uno verso l'altro con le pendenze che li connotano, daranno linee di estrapolazione che si sovrappongono o quasi (fig 1.3). In realtà tale relazione spaziale va verificata nelle tre dimensioni e non in due sole, per cui si deve lavorare almeno su set di profili diversamente orientati o, meglio, su rappresentazioni ed estrapolazioni tridimensionali. Inoltre le estrapolazioni di cui sopra non possono essere lineari



*Fig 1.1 Sia nei terrazzi deposizionali (A) che su roccia (B) si possono evidenziare livelli disaccoppiati (da Merritts et alii 1994, modificato)*

se non su brevi distanze, poiché le antiche superfici con le quali si tratta quasi mai erano esattamente planari su lunghe distanze, ma quasi sempre leggermente concave ed asintoticamente tangenti verso valle al livello di base che ne controllò la genesi.

La delineazione dei confini dei terrazzi in un sistema a più ordini, può essere effettuata mediante un'analisi speditiva (sul campo, foto aeree e/o carte topografiche) e risulta piuttosto agevole quando non vi sono forti rimodellamenti tardivi delle superfici (per erosione e/o deposizione) né una riduzione a crinali per fitta dissezione e forte smussamento degli orli. Si tratterà di individuare e mappare il margine interno (limite inferiore della scarpata fluviale di un ordine di



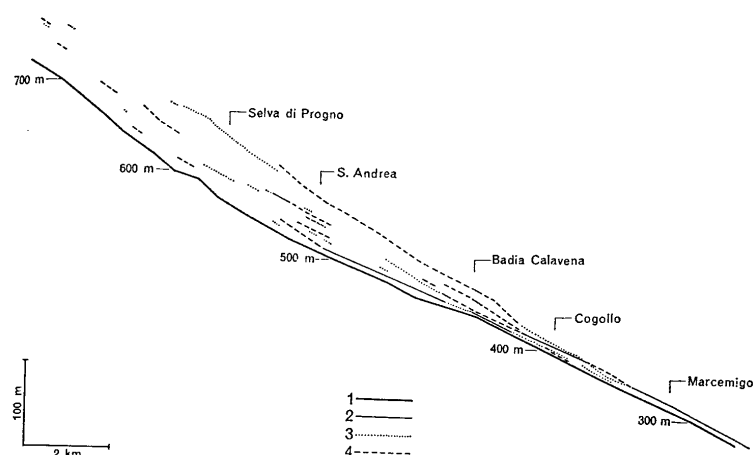
*Fig 1.2 Generazione di livelli di terrazzi fluviali deposizionali su più ordini. (A) Livelli erosionali t2, livelli deposizionali t1 (B) Livelli deposizionali non coevi t1 e t2 (da Castiglioni, 1986)*

terrazzi più antico) ed il margine esterno (ciglio della scarpata del terrazzo medesimo). Ma questi elementi possono essere facilmente individuati, come vedremo, anche in automatico da DTM tramite l'estrazione della mappa delle pendenze e la selezione dei valori tipici delle rotture di pendenza caratteristici delle scarpate fluviali. Ciò non risulta però sufficiente ad identificare gli ordini, ma semmai ad isolare singoli lembi di cui rimane ignota la correlazione con altre superfici. Quindi per vagliare la correlabilità dei lembi e proporre una suddivisione in più ordini, ove ciò sia imposto dalla realtà geomorfologia (paesaggi policiclici), sono necessarie ulteriori elaborazioni che considerino altri attributi oltre che la sola pendenza.

Sia che si tratti di terrazzi di origine deposizionale o erosionale, l'area di studio deve essere analizzata utilizzando come livello guida iniziale il fondovalle,



il quale fornisce i primi essenziali vincoli circa le direzioni e le entità di pendenza del sistema vallivo, almeno durante gli ultimi stadi della sua evoluzione. Il fondovalle viene ad essere quindi identificato come livello di base rispetto agli ordini posti a quote più elevate. Detto in tali termini, l'analisi e la identificazione degli ordini di terrazzi sembra essere un compito abbastanza semplice. Tuttavia all'atto pratico si rileva assai arduo riuscire ad identificare i vari ordini per una ragione di tipo strettamente geometrico. In effetti è noto che il livello di base locale, per definizione, segue il percorso sia planare che altimetrico del corso d'acqua principale e si riflette con progressivi incrementi (crescenti con leggi variabili nel tempo e nello spazio) nelle varie valli tributarie. In letteratura sono note diverse funzioni che tentano di descrivere con legge geometrica il tipico andamento del profilo longitudinale dei corsi d'acqua. Nei lavori di Hovius (2000) si cerca di individuare una correlazione statistica tra profili del fondovalle di differenti bacini della stessa area di studio, utilizzando funzioni polinomiali per descriverne la forma. Ma tali formule si adattano semmai a corsi di alto ordine e molto maturi che caratterizzano regioni di antica stabilità geologica e lunga storia morfoevolutiva. Più spesso il profilo longitudinale di un corso d'acqua, specie se di zone geologicamente giovani come quella appenninica, devia sensibilmente dal caso di cui sopra, risultando composto di più elementi (profili discontinui) tra cui addirittura alcuni convessi, insieme a salti (nickpoint) che separano vari tratti complessivamente concavi e con gradi di concavità solo mediamente decrescenti da monte verso valle. Tali salti possono essere il risultato di un'onda di erosione regressiva.



*Fig 1.3 Profili longitudinali di lembi di terrazzi fluviali confrontati con il profilo longitudinale dell'alveo attuale (da Castiglioni 1986).*

Simili profili compositi sono legati a variazioni di litologia, a disturbi tettonici recenti o ancora attivi, a variazioni glacio-eustatiche o alle variazioni locali (a scala di sottobacino) della tessitura degli apporti detritici dai versanti e del rapporto tra carico limite e carico reale.

La descritta variabilità del profilo longitudinale si riflette anche nella distribuzione dei gradienti dei terrazzi, la cui disposizione è infatti controllata dai gradienti e dalle variazioni di gradiente che si avevano lungo un antico fondovalle. Si possono così individuare ordini di terrazzi caratterizzati da vergenze longitudinali differenti rispetto al fondovalle. Alcuni ordini possono risultare convergenti mentre altri divergere verso valle (Castiglioni 1986). Ciò che complica ulteriormente l'analisi, è che quando si proietta su di un grafico altezze-distanze un profilo longitudinale, si introduce spesso la semplificazione di non seguire tutte le curve (talora anche ignote) che l'antico fondovalle seguiva nel tratto considerato. Come già detto, i sistemi vallivi ed i terrazzi fluviali sono entità tridimensionali che come tali andrebbero sempre trattate ed analizzate. Infatti anche lembi di terrazzi appartenenti alla stessa fase di formazione, possono avere vergenze e pendenze variabili da zona a zona del sistema vallivo gerarchizzato cui appartengono (fig 1.3 ed 1.4).

Un metodo classico con cui è possibile individuare gli ordini dei terrazzi

fluviali, siano essi erosionali che deposizionali è lo studio di profili topografici (cfr. capitolo 2). I livelli, individuabili in maniera locale tramite appunto l'analisi di profili sia trasversali che longitudinali, risultano spesso essere di difficile attribuzione a causa della non linearità della distribuzione delle quote sia lungo il corso principale sia lungo i tributari. Inoltre l'utilità dell'analisi dei singoli profili spesso risulta nella sola possibilità di individuare un ordinamento locale dei terrazzi. Per poter ricavare un ordinamento che sia valido per l'intero percorso del fondovalle o addirittura per l'intero sistema vallivo (ad esempio un ordinamento a livello di intero bacino idrografico) risulta necessario analizzare

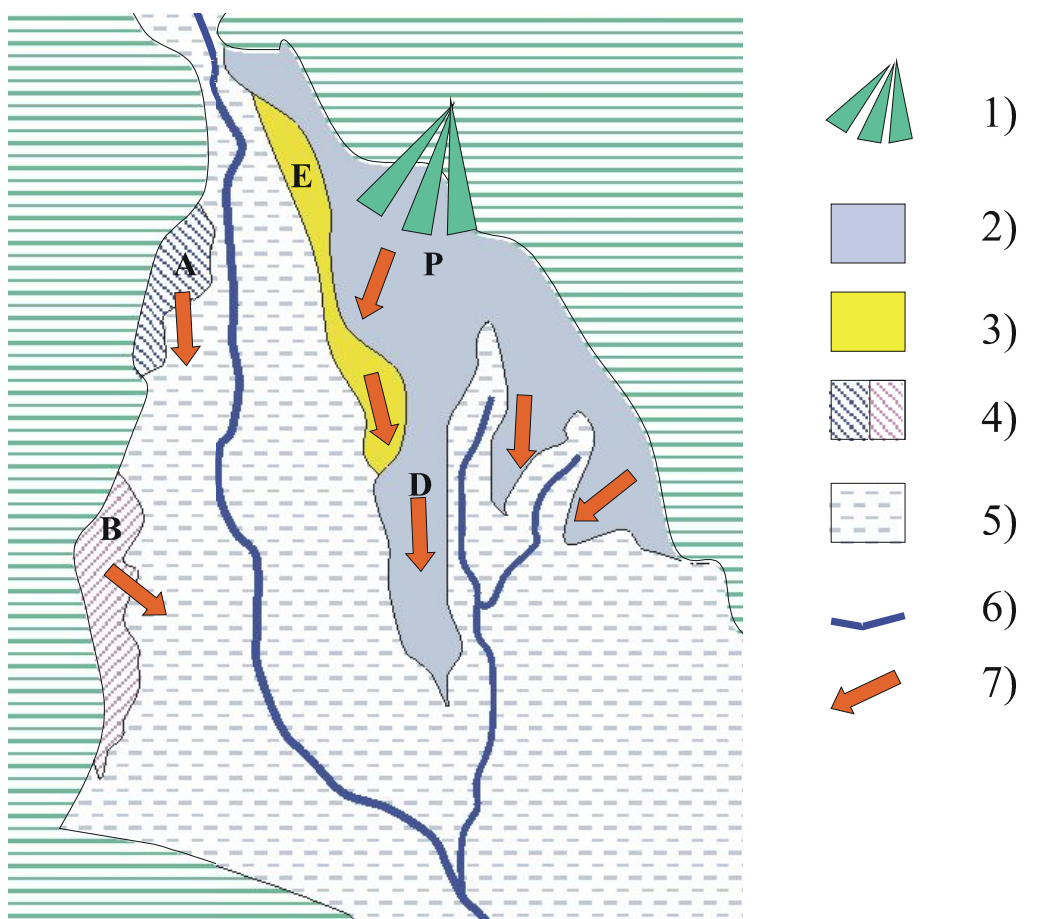


Fig 1.4 Schema planimetrico (teorico) della morfologia associata a terrazzi fluviali in scenari orogenici  
 1) conoidi 2) Lembi di terrazzi di I ordine 3) Lembi di terrazzi di II ordine 4) lembi di difficile attribuzione 5) fondovalle 6) Reticolo idrografico 7) Possibili vergenze dei lembi di terrazzo

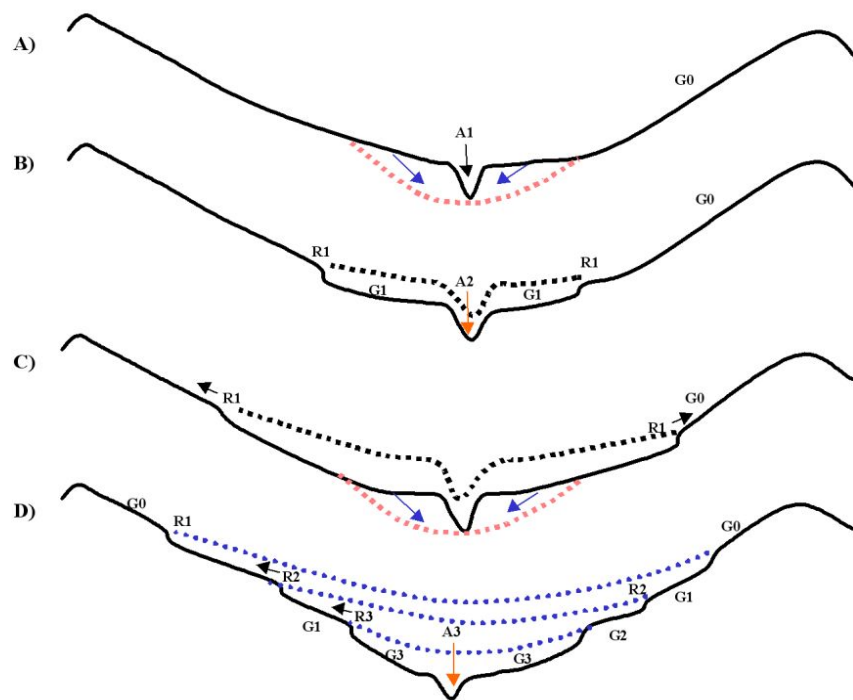
sia i profili trasversali che i profili longitudinali. Dai profili trasversali è possibile ottenere un ordinamento dei profili longitudinali collegando tramite estrapolazione i vari segmenti ottenuti. Nel caso lo scopo sia quello di individuare i vari ordini lungo un singolo corso d'acqua, può essere sufficiente l'analisi dei soli profili longitudinali estratti tenendo conto del percorso planimetrico degli stessi. La problematica insita in questo tipo di analisi risulta evidente nei casi di terrazzi erosi. L'erosione è dovuta principalmente all'evoluzione dei tributari del fiume che occupa il fondovalle. In tal caso spesso, soprattutto nel caso di piane alluvionali evolutesi successivamente a fasi di approfondimento del reticolo idrografico, si possono rinvenire terrazzi sotto forma di lembi a volte notevolmente distanti anche se appartenenti alla stessa fase di deposizione o di erosione (lembi A e B di fig 1.4). Il paesaggio in tali casi risulta come una valle dai fianchi acclivi e variamente dissecati, ed un fondovalle in cui sono riconoscibili una piana alluvionale e terrazzi di vario ordine ma la cui correlazione non risulta evidente dall'analisi della semplice carta topografica. Ciò è dovuto spesso al fatto che la combinazione tra bassa pendenza delle forme, scarpate che separano vari ordini oramai degradate da processi di erosione quali fenomeni di scalzamento al piede, la presenza di conoidi che possono obliterare le forme, fanno sì che l'andamento delle isoipse possa ingannare l'occhio e che i profili topografici non seguano più un andamento regolare rettilineo, ma irregolare, e quindi risulta più arbitrario il criterio con cui si tracciano le rette che collegano i vari lembi. La distribuzione non costante lungo il suo percorso, dell'energia del corso d'acqua, generalmente dà origine ad un sistema di terrazzi il cui andamento altimetrico rispecchia la variabilità del gradiente originale del corso d'acqua al momento della formazione del terrazzo di tipo deposizionale. Nel caso di corsi d'acqua che presentano energia molto bassa e quindi gradiente altimetrico molto modesto (misurabile in in decimetri per chilometro) risulta evidente che anche i livelli ad esso riconducibili risulteranno caratterizzati da una quota pochissimo variabile (e variabile con maggiore regolarità) lungo tutto il

piano di sviluppo del sistema fluviale. Si tratta –come già accennato- di casi ristretti a corsi d'acqua molto maturi o a tratti finali di corsi di media maturità nei quali , una volta determinata in un punto l'altezza relativa di un certo ordine di terrazzi , si possono cercare lembi correlabili più a valle e più a monte (anche se ci si sposta di chilometri) utilizzando lo stesso valore di dislivello o comunque regole monotone e semplici di incremento/decremento.

## **1.2 Glacis d'erosione in rocce tenere**

I Glacis d'erosione in rocce tenere (Glacis d'erosion en roche tendre, [Biot et alii 1958, Vaumaz 1963] ), sono superfici relitte di origine erosionale, resti di spianamenti prodottesi durante fasi di svasamento di una valle. Il processo attraverso cui si formano queste superfici a bassa pendenza, sospese spesso a diversi metri di altezza rispetto all'attuale fondovalle, è la combinazione di una fase di addolcimento dei fianchi vallivi, protrattasi durante fasi di maggiore stabilità da parte del reticolo idrografico, seguita da fasi di maggiore tendenza all'approfondimento ed alla dissezione delle forme sia di erosione che di accumulo. Spesso tali superfici relitte sono, analogamente ai terrazzi fluviali, presenti con più ordini di generazioni. Quando al piede del versante è attivo un processo di trasporto, si forma un versante di *replacement* che può assumere diverse conformazioni seconda dei processi dominanti e delle proprietà litologiche del substrato. In questo caso è necessario dare una possibile spiegazione a come una evoluzione verso un versante policiclico, ossia caratterizzato da più ripetizioni della sequenza descritta all'inizio, possa determinare la formazione e la conservazione di più ordini di tali superfici. Il processo determinate che si suppone sia alla base della formazione di tali superfici è la recessione del versante per *slope replacement* lungo versanti con un angolo di pendenza moderato (Castiglioni 1986).

Contrariamente ai versanti di recessione impostati su substrato coerente e ben stratificato quali calcari, ed in condizioni morfoclimatiche prevalentemente di tipo semi-arido, l'evoluzione non prosegue secondo la modalità descritte da Lehmann (1933). Infatti sia la natura del substrato sul quale sono impostati questi glacis, sia le condizioni morfoclimatiche sono differenti. Le oscillazioni del clima, da periodi caratterizzati da forti precipitazioni ad un clima più di tipo periglaciale, hanno fatto sì che in differenti momenti dell'evoluzione del sistema vallivo alcuni processi prevalessero su altri. In particolare, data la minor coerenza e differente comportamento del substrato (prevalentemente flyscoide nel caso delle aree test analizzate), è possibile che il mass-wasting, durante le fasi di svasamento, abbia avuto un ruolo determinante nel rimodellare le superfici conferendogli quelle pendenze molto più basse rispetto alle pendenze caratteristiche dei versanti di Lehman. Lo schema di figura mostra tale possibile spiegazione della policiclicità dei versanti a glacis. In fig 1.5a dopo un'iniziale fase di approfondimento del reticolo durante un periodo di clima caldo-umido, si passa ad un periodo caratterizzato da bassa piovosità. A causa della minore energia dei corsi d'acqua, all'approfondimento prevale il dilavamento e la recessione delle scarpate fluviali, che determina la degradazione delle superfici esistenti con un'arretramento dei versanti attuali e la formazione di nuovi posti a quote inferiori (fig 1.5b). Un successivo ciclo, caratterizzato dall'approfondimento del reticolo idrografico, porta alla dissezione del versante, formando lembi isolati. I glacis possono poi essere ridotti a creste o rimodellati da fenomeni franosi.



*Fig 1.5 Schema di possibile formazione di Glacis vallivi sospesi G0,G1,G2,G3 glacis; R1,R2,R2 cornici sommitali; A1 A2 A3 fasi di approfondimento dell'alveo..le linee blu rappresentano la presunta paleovalle ottenibile in linea teorica tramite un processo di estrapolazione delle superfici, le frecce blu le fasi di dissezione da parte dei tributari.*

### 1.3 Altre superfici relitte

Altre superfici relitte sono il risultato di planazioni avvenute in epoche più remote. I più noti modelli evolutivi del rilievo, di Davis (1899), King (1976), e Penck (1953), nonostante le differenze nei rapporti temporali tra agenti endogeni e processi esogeni, propongono una spiegazione dell'origine di tali ampie superfici. Il modello di Davis, ampiamente criticato e concettualmente non più attuale, descrive un'evoluzione del rilievo caratterizzata da episodiche fasi di sollevamento tettonico seguite da lunghi periodi di stabilità durante i quali, a causa principalmente dell'azione dei processi fluvio-denudazionali, si raggiunge un rilievo a bassissima energia, quasi planare, definito dall'autore come penepiano. Tale fase dell'evoluzione del rilievo viene definita fase di maturità. Una eventuale nuova (e breve nel tempo secondo l'autore) fase di sollevamento

porta alla ripresa dell'approfondimento del reticolo idrografico e quindi ad un "ringiovanimento del rilievo". L'evoluzione procede quindi per cicli che possono portare anche ad un tipo policiclico in cui si riconoscono relitti del precedente ciclo.

Il modello di Penck si discosta da questa interpretazione ciclica ipotizzando una graduale ma costante presenza della tettonica. La demolizione del rilievo avviene quindi contemporaneamente al sollevamento. L'autore ipotizza un'evoluzione del rilievo per *slope replacement*. Se prevale il sollevamento si genera un paesaggio caratterizzato da versanti convessi. Con il declino del ritmo di sollevamento si giunge alla formazione di un rilievo a bassa pendenza dalla forma concava definito *endrumpf*.

Il modello di King della pediplanazione condivide con quello di Davis la episodicità del sollevamento ma si discosta per quanto riguarda i processi che portano alla planazione. Il versante subisce un arretramento parallelo lasciandosi dietro un rilievo planare definito dall'autore appunto pedipiano. Si riconosce in questo modello evolutivo la presenza di un'evoluzione verso paesaggi caratterizzati da pedipiani di età differente correlabile con la quota. L'evoluzione del pedipiano continua anche dopo la fase di sollevamento, per tutte le generazioni di superfici, che al limite possono essere totalmente cancellate dai processi di arretramento del versante.

Al di là dei modelli interpretativi sopra descritti e delle relative terminologie, tali paesaggi prendono il nome generico di superfici di spianamento (Castiglioni 1986). Si tratta di superfici che rasano indifferentemente rocce di vario tipo in discordanza con la stratificazione. Se il livello di base cambia, per cause eustatiche e/o tettoniche, vi è una ripresa dei processi di approfondimento del reticolo e dei processi denudazionali di versante e conseguente aumento della pendenza lungo gli stessi. La conservazione di superfici isolate a basso angolo di pendenza può essere dovuta a vari fattori, quali ad esempio l'erosione selettiva. Tali superfici relitte rimangono sospese su di un livello che rispecchia l'antico



livello di base prima dell'approfondimento del reticolo.

Un esempio molto diffuso di superfici relitte che si rinvencono lungo i fianchi vallivi dell'Appennino Centro-Meridionale sono le cosiddette “Paleosuperfici” (Dainelli 1926). Queste hanno un'articolazione altimetrica molto variabile e scarsa continuità laterale tanto che si presentano spesso come lembi sparsi di piccole dimensioni. La presenza di più ordini di generazioni, dovuta a fasi pol cicliche nell'evoluzione dei sistemi vallivi, se da un lato rende plausibile il tentativo di una ricostruzione spazio temporale dell'evoluzione di tali sistemi, dall'altro implica complesse procedure di valutazione che devono tenere conto di innumerevoli parametri. Difatti non è semplice effettuare correlazioni spazio temporali sia perché, essendo superfici erosionali, spesso non è possibile rinvenire lungo tali forme depositi attribuibili alla originaria superficie di spianamento, sia appunto per la scarsa continuità laterale delle stesse (Aucelli et alii 2003, Amato 1995). Essendo ancora più antiche delle superfici già descritte, i processi endogeni hanno avuto più tempo per modificare l'originale assetto altimetrico delle stesse. Spesso lembi relitti di piccola estensione, non contigui lateralmente, risultano non raccordabili altimetricamente anche se la loro vicinanza geografica porterebbe a raggrupparli in un unico lembo. Si può ipotizzare in tali casi che la superficie originaria sia stata smembrata dalla tettonica ed i lembi dislocati siano stati rimodellati da processi di versante che possono averne ridotto sia la superficie che la quota.

*Parte B*  
*Dati di letteratura*

## **B.I Concetti di base**

### **2.1 Il GIS**

I GIS o Sistemi Informativi Geografici (Wilson & Gallant 2000, O'Sullivan & Unwin 2003) racchiudono un'insieme di funzionalità che li rendono strumento informatico di enorme utilità nell'analisi spaziale. Tuttavia, in campo geomorfologico il GIS spesso non rivela immediatamente le sue potenzialità, specialmente quale strumento di supporto alle analisi delle forme ed allo studio dell'evoluzione del paesaggio.

GIS è il termine col quale ci si riferisce comunemente al pacchetto software in grado di gestire dati geografici. In realtà è più correttamente un insieme di risorse umane, informative ed hardware. Il sistema software/hardware propriamente detto, è un sistema in grado di acquisire ed archiviare dati geografici, manipolarli, restituire il dato tramite mappe (a video o tramite stampati) o tabelle. Rispetto ai sistemi classici di archiviazione elettronica, il *database* di un GIS si distingue per l'associazione dei *record* a delle coordinate spaziali. La tecnologia dei sistemi GIS è relativamente nuova ed ampiamente utilizzata solo con lo sviluppo dei personal computer, e quindi dei software GIS, a partire dagli anni 80. La chiave del loro valore enorme è che offrono agli utenti l'opportunità di analizzare e manipolare *database* di grandi dimensioni con gli strumenti di analisi più svariati, l'utilizzo di tematismi, la ricerca tramite "query" per caratteristiche particolari in aree particolari tramite sintassi standard (SQL), l'aggiornamento rapido dei *database* e la possibilità di collegamento con database esterni (che possono essere relazionali od ad oggetti) in maniera da poter condividere i dati con altri sistemi software. Altre caratteristiche peculiari sono la possibilità di poter essere impiegati per la modellizzazione di fenomeni spaziali e più recentemente realizzare sistemi interattivi utilizzabili via web in grado di

restituire, tramite interfacce facilmente utilizzabili anche da utenti non esperti, l'informazione geografica contenuta nei *database* di un server: i cosiddetti *web-gis*.

I dati utilizzabili in un GIS possono essere divisi in due categorie di modelli: vettoriali e *raster*. In breve un dato vettoriale è rappresentato da punti, linee, e poligoni, mentre un dato del *raster* è rappresentato da *pixel* (cfr. capitolo 3). Entrambi i modelli posseggono i loro vantaggi e svantaggi. Il modello *raster* è configurato come una struttura dati molto semplice e facilmente manipolabile, la struttura vettoriale può essere complessa ma restituisce una rappresentazione accurata di confini e tematismi lineari e puntuali. Si discuterà più estesamente di tali differenze, in particolare sulle caratteristiche del dato *raster*. In particolare il dato *raster* si presta bene alla manipolazione dei modelli digitali del terreno. Il pacchetto GIS utilizzato durante la tesi di dottorato è un gis open-source: GRASS versione 5 installato su sistema operativo Linux. Si tratta di un GIS che, oltre ad essere open-source, si caratterizza dalla grande flessibilità e numero di funzionalità presenti.

## **2.2 Approccio metodologico**

L'analisi geomorfologica di cui si tratta in questa tesi si avvale di strumenti informatici e delle conoscenze derivanti da varie discipline: scienze della terra, matematica, statistica, ingegneria ecc... Difatti, come sottolinea Pike (1995), la geomorfologia quantitativa si è evoluta e continua ad evolvere rapidamente, sia grazie alla sempre maggiore disponibilità di tecniche per l'acquisizione di modelli digitali del terreno che allo sviluppo della potenza di calcolo dei personal computer.

## 2.3 La geomorfologia quantitativa

Gli studi sulla descrizione quantitativa delle caratteristiche topografiche della superficie terrestre, rispetto all'analisi puramente qualitativa, hanno portato a definire e a proporre tutta una serie di metodologie e parametri utili ad una descrizione di tipo oggettivo delle caratteristiche della superficie terrestre. Horton (1945) è il pioniere della analisi quantitativa del reticolo idrografico; anche se solo dal punto planimetrico. Analisi del profilo del versante e degli angoli di pendenza caratteristici delle regioni di studio (Young 1972), Ahnert utilizza il concetto di local relief (Ahnert 1984) al fine di individuare correlazione tra estensione e quota massima delle catene montuose, Strahler (1957) introduce il concetto di integrale e curva ipsometrica quale strumenti oramai consolidati di analisi dei bacini fluviali. Hovius (2000) utilizza la regressione di tipo logaritmico e le derivate di funzioni polinomiali al fine di individuare relazioni comuni nel profilo longitudinale di numerosi fiumi nella regione di Taiwan utilizzando anche le derivate per individuare relazioni comuni.. Di recente anche la dimensione frattale introdotta da Mandelbrot (1982) ha fatto il suo ingresso nella analisi e caratterizzazione della superficie terrestre in contrapposizione al concetto di superficie continua (Outcalt et alii 1994, Goodchild & Mark 1987, Wood 1996, Zahouani et alii 1998). Oltre alla semplice analisi dei cosiddetti “parametri morfometrici” (cfr. capitolo 3), che descrivono le proprietà geometriche locali della superficie (Evans 1972), è possibile effettuare analisi e classificazioni della superficie per unità omogenee, sia tramite metodi manuali (Hammond 1964a, Speight 1967) che automatici (Dikau 1995). Tale approccio è stato perlopiù utilizzato in campi applicativi, quali l'analisi dei rischi geoambientali, associando analisi di tipo combinatorio con tecniche di tipo statistico (Carrara et alii 1995). Uno dei problemi spinosi dell'approccio quantitativo alla geomorfologia attraverso i modelli digitali del terreno, risiede nella forte dipendenza dei risultati ottenibili cambiando la scala di analisi.

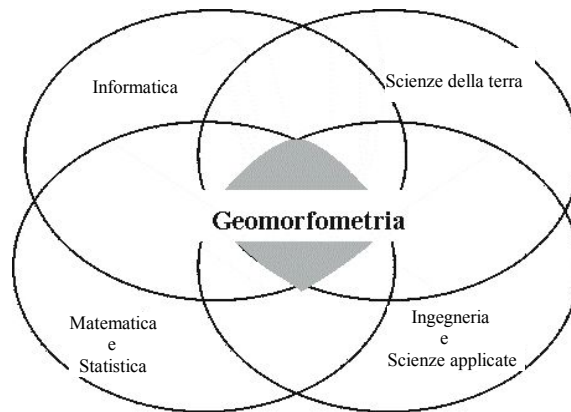
Diminuendo la risoluzione si perdono infatti informazioni relative alla geometria della superficie, ossia l'andamento della superficie a partire dal cosiddetto microrilievo. Nel caso il tipo di studio si basi sull'analisi del cosiddetto macrorilievo, risulta utile poter utilizzare un modello di superficie che evidenzii solo i cambiamenti a piccola scala. La distinzione tra macrorilievo e microrilievo è legata alla lunghezza d'onda alle quali si presentano le variazioni altimetriche, o irregolarità, della superficie (Stone & Dungundji 1965). Dal punto di vista geomorfologico rispecchiano processi che agiscono anch'essi a scale differenti ed evidentemente anche lungo archi di tempo differenti. Il microrilievo scaturisce da processi quali soliflusso, *rill erosion*, *gull erosion*, *wethering* ecc.. Il macrorilievo si riferisce ad elementi che si presentano a lunghezze d'onda maggiori e sono legati a processi quali ad esempio la tettonica o i rimodellamenti del versante quali il *parallel retreat*. Ovviamente i termini suddetti sono solamente indicativi poiché la superficie topografica è caratterizzata da elementi che si presentano a scala e quindi lunghezza d'onda variabile: dai terrazzetti prodotti dal soliflusso alle grandi catene montuose (megarilievo). Per questo, al fine di una caratterizzazione completa del rilievo risulta spesso necessario evidenziare sia gli elementi di grandi dimensioni che le variazioni locali della superficie. Alcune tecniche di analisi automatica sono state proposte per risolvere il problema dell'analisi di tipo multiscala. Esistono strumenti matematici che permettono un'analisi multiscala dei modelli digitali del terreno, e che si basano sull'analisi armonica o l'uso di filtri wavelet (Zatelli & Antonello 2002, Callant & Hutchinson 1997, Harrisont & Lo 1996) ma un'analisi degli elementi che si presentano a scala più fine non può prescindere dall'utilizzo di modelli digitali ad alta risoluzione. La disponibilità gratuita di DEM soprattutto a scala globale, quali il GTOPO30 (cfr. paragrafo 3.3), ha dato impulso a numerosi studi di geomorfologia tettonica (Summerfield 2000, Jordan 2000, Florinsky 1996) che interpreta l'evoluzione del paesaggio attraverso l'ausilio dell'analisi dei DTM. I DTM a risoluzione più elevata, raramente disponibili gratuitamente, sono stati

impiegati estensivamente nel campo della geomorfologia dinamica tramite automi cellulari, per l'analisi di modelli idraulici, processi di erosione e sedimentazione ed evoluzione del versante (Coulthard et alii 1997, Mitas et alii 1996, Luo 2001) e per estendere l'analisi del reticolo idrografico oltre il contesto del singolo bacino (Verosmarty et alii 2000).

Una lista estesa delle numerose pubblicazioni sull'argomento attinente alla disciplina descritta nel paragrafo successivo e definita con il termine di geomorfometria, ed in genere sulle numerose pubblicazioni di geomorfologia quantitativa ed applicazioni dei modelli digitali del terreno e dei GIS in campo geomorfologico, è possibile trovarla in (Pike 2002). Le numerose sottotematiche ivi citate rispondono ad esigenze di carattere immediato e sono state oggetto pertanto sia in anni recenti che in anni passati di numerosi studi che hanno portato alla formalizzazione di metodologie e alla loro implementazione all'interno di pacchetti GIS sotto forma di funzionalità (*tools*) del pacchetto stesso. L'analisi dell'evoluzione di medio e lungo termine del paesaggio risponde ad esigenze diverse, proprie della geomorfologia classica. Tali analisi non possono essere affrontate senza prima caratterizzare in maniera approfondita la struttura del paesaggio ad esempio tramite la cartografazione di forme specifiche del rilievo.

## **2.4 Geomorfometria**

Il concetto di “geomorfometria” come scienza a se state, emersa come incontro di differenti discipline, quali quelle menzionate sopra, è un concetto recente che però, vista la vastità delle conoscenze coinvolte, può destare qualche perplessità dal punto di vista di quale si in effetti l'oggetto di studio e le finalità della disciplina. Basta guardare il diagramma di fig 2.1 per rendersi conto delle interconnessioni tra studio, discipline e finalità. Pike richiama il concetto di morfometria come sinonimo di geomorfometria. Non sembra tuttavia chiaro



*Fig 2.1 Origine della geomorfometria come scienza derivata dall'integrazione di diverse discipline (da Pike 1995, modificato)*

come i due termini possano riferirsi esattamente allo stesso oggetto di studio. Infatti la morfometria richiama lo studio della forma in quanto tale, senza nessuno specifico riferimento alla natura della stessa. Tuttavia, dal punto di vista degli strumenti di analisi, non vi è differenza nell'analizzare una porzione di superficie terrestre o una superficie di altro tipo quando, come spesso avviene, ciò che si analizza è solo un modello della realtà. Se il modello che descrive i vari oggetti è equivalente dal punto di vista della struttura, possono essere identiche anche le metodologie utilizzate per analizzarlo. Per esempio la cosiddetta metrologia (Pike 2001), scienza che si occupa dello studio delle superfici a scala micrometrica e nanometrica, utilizza metodologie di analisi simili a quelle proposte ed impiegate per l'analisi dei DTM. Lo stesso vale per le varie tecniche di analisi delle immagini digitali. Termini quali spartiacque, segmentazione, gradiente, lineamenti ecc.. sono comunemente riscontrabili nelle pubblicazioni delle summenzionate discipline. Per quanto riguarda il termine geomorfometria si intende qui evidentemente lo studio della forma della superficie terrestre piuttosto che una superficie qualsiasi e di fatto l'indirizzo morfometrico della geomorfologia.

Evans (1972) sottolinea come la morfometria, applicata alla geomorfologia, può essere distinta in morfometria generale e morfometria specifica. La



morfometria specifica è mirata allo studio di forme discrete, ovvero strutture e caratteristiche della superficie distinguibili quali elementi caratteristici del paesaggio, come ad esempio il reticolo idrografico, vulcani, scarpate, terrazzi fluviali ecc.. Ovvero ciò che può essere descritto anche da un punto di vista qualitativo ma che la geomorfometria, tramite l'approccio quantitativo, è in grado di descrivere in maniera meno soggettiva. La geomorfometria generale invece si occupa dell'analisi della superficie terrestre attraverso lo studio delle sue proprietà in quanto superficie continua, anche se in effetti la superficie terrestre non è realmente continua e differenti approcci, quali l'analisi frattale (Outcalt et alii 1994), sono stati applicati all'analisi della superficie terrestre. In ciò la geomorfometria generale è utile a definire la forma lungo la superficie in quanto tale senza entrare nel merito della natura della stessa. Tuttavia l'individuazione automatica di forme specifiche non può prescindere dalla analisi delle caratteristiche delle proprietà della superficie continua poiché da queste derivano i parametri necessari a delimitare e classificare le forme specifiche.

Negli ultimi anni l'indirizzo morfometrico è stato caratterizzato dalla sempre maggiore presenza di tecniche di analisi che si avvalgono dell'elaborazione elettronica. Grazie a questo fatto, si è ampliata anche il numero di applicazioni, ad esempio nel campo dell'idrologia, dei rischi geoambientali, della cartografazione tematica, della tettonica, della batimetria ecc.. prima confinati ad analisi puramente qualitative a causa della complessità di un approccio quantitativo attraverso i metodi tradizionali. La disponibilità di computer potenti ed a basso costo ha in particolare permesso lo sviluppo di strumenti informatici specifici per lo studio dei modelli digitali del terreno (DEM e DTM) che permettono di quantificare e visualizzare le caratteristiche della superficie terrestre per vaste aree. Al giorno d'oggi numerose procedure per il calcolo dei cosiddetti parametri morfometrici sono direttamente implementate all'interno di pacchetti GIS commerciali ed in software specialistici. La disponibilità e la qualità dei dati topografici costituisce la principale problematica dell'analisi

automatica della superficie terrestre. Tuttavia non basta la disponibilità di dati di buona qualità per poter affrontare in maniera concisa e disambigua lo studio della forma della superficie terrestre.

## **2.5 Metodi di analisi tradizionale (quantitativa e qualitativa)**

La cartografazione delle forme affrontato attraverso metodi classici, prevede l'analisi estensiva attraverso l'uso di mappe topografiche a varie scale, foto aeree ecc. L'analisi tradizionale si caratterizza per il notevole tempo necessario alla lettura delle fonti ed alla redazione degli elaborati cartacei a scala regionale.

La geomorfologia morfogenetica in particolare distingue le forme a seconda della loro origine e quindi dei processi che le hanno modellate. La sola individuazione e distinzione delle forme, attraverso i metodi classici di analisi, comporta principalmente lo studio di carte topografiche, foto aeree, ed il rilevamento sul campo. Se il contesto di studio è a scala regionale è evidente che queste operazioni comportano un notevole dispendio di tempo. Ma le operazioni menzionate sarebbero inutili senza l'interpretazione delle forme. Ciò comporta l'analisi approfondita della natura del paesaggio e delle sue caratteristiche, analisi che devono essere appropriate al tipo di studio geomorfologico che si intende svolgere. La complessità della struttura, sia a scala regionale che a scala locale, e la grande quantità di dati necessari allo studio dei fenomeni morfoevolutivi rende necessario l'utilizzo di approcci alternativi e/o complementari agli studi di tipo tradizionale.

### **2.5.1 Analisi di foto aeree**

Diffusa e molto utilizzata quando sono disponibili le coperture dell'area da studiare. Le foto aeree sono strumenti che consentono, sia grazie all'elevata risoluzione che la pellicola riesce a garantire, ma soprattutto grazie al fatto di

poter sfruttare l'effetto stereoscopico, una visione dettagliata delle forme ed anche, in molti casi, di poter riconoscere negli affioramenti rocciosi, la stratificazione. Tramite foto aeree è spesso possibile distinguere le diverse litologie affioranti ed i contatti tettonici. Ciò le rende indubbiamente lo strumento ideale per individuare elementi quali scarpate, forme di origine carsica ecc.. Vengono studiate sia per realizzare una cartografia estensiva, che per analizzare singoli fenomeni di interesse. Il loro utilizzo si basa molto sull'analisi soggettiva delle caratteristiche dell'area di studio. Difatti la forma della superficie, grazie all'effetto stereoscopico, risulta ben evidente, ma l'analisi dei vari elementi del paesaggio è possibile solo da un punto di vista qualitativo visto che risulta impossibile quantificare i parametri, quali angoli di pendenza, concavità e convessità delle forme con l'ausilio di un normale stereoscopio. Mentre si presta bene, come già accennato, ad individuare singoli elementi quali scarpate di faglia, piane alluvionali ecc.. e risulta quindi adeguata alla interpretazione della genesi del paesaggio.

### **2.5.2 Carte topografiche**

Lo strumento tradizionale più adatto all'analisi quantitativa è la carta topografica. Il dettaglio con cui è restituita la superficie topografica dipende dalla scala della carta utilizzata e conseguentemente dalla equidistanza delle isoipse. Rispetto alle foto aeree, tramite la carta topografica è possibile avvalersi di metodi di analisi quantitativa. Dalle isoipse è possibile ad esempio estrarre profili topografici e studiarli sia nella loro forma da un punto di vista qualitativo, che estrarne i relativi parametri espressi quantitativamente. In tal modo è possibile effettuare analisi statistiche di tali parametri (quota, pendenza, curvatura ecc..) riferiti ad intere aree di studio. E' possibile ad esempio una volta delineato un bacino idrografico, calcolarne l'area od altri parametri derivati quali curve ipsometriche, pendenza media ecc..

### **2.5.3 Analisi del profilo topografico**

Una trattazione dei metodi di analisi tradizionale non può prescindere da una descrizione delle metodologie di analisi del profilo topografico. La forma del profilo è infatti una caratteristica topografica di fondamentale importanza per lo studio ed individuazione dei processi attivi e di quelli passati che hanno generato una determinata forma del rilievo. L'analisi quantitativa è certamente lo scopo principale dell'analisi del profilo. Tuttavia nonostante la sua importanza per l'analisi delle forme, tradizionalmente non è stata sempre sfruttata come ausilio alla cartografia geomorfologica, quanto piuttosto per l'analisi dei processi di versante, in particolare quelli legati al trasporto fluviale. Per quanto riguarda le metodologie di studio tradizionali, si possono qui elencare alcune delle metodologie maggiormente utilizzate per la realizzazione di carte di sintesi delle caratteristiche geomorfologiche di un'area di studio. La prima, la foto aerea, non risulta adatta all'analisi del profilo, almeno fintanto che si utilizza come strumento che non consente misurazioni, quale lo stereoscopio. Dalla seconda, la carta topografica, è invece possibile estrarre il profilo e studiarlo quindi nella sua forma.

Young (1972) descrive come l'analisi della forma del rilievo sia possibile attraverso l'individuazione di varie caratteristiche e tipologie del profilo topografico. Difatti, come in effetti è facilmente dimostrabile da una descrizione analitica della superficie (cfr. capitolo 3), questo può essere analizzato quale profilo longitudinale, ossia la forma del versante lungo la direzione di massima pendenza, o quale profilo planare, ossia la forma del profilo del versante valutata seguendo il percorso delle isoipse e che conseguentemente descrive la curvatura delle linee di livello. La forma longitudinale può essere analizzata sia in termini di concavità e convessità che descrivendo gli angoli di pendenza lungo lo stesso. La pendenza viene tradizionalmente misurata in percentuale, ovvero il rapporto tra differenza di quota tra due punti e la distanza planimetrica moltiplicato un fattore cento, od in gradi calcolando l'arcotangente di tale rapporto.

Analizzando l'insieme di queste fondamentali proprietà del profilo è possibile trarne una caratterizzazione quantitativa.

Il profilo può essere scomposto in una successione di segmenti in base alle caratteristiche di pendenza e concavità. Ogni singola parte, definita come elemento da Young (1972), è caratterizzata da differenti combinazioni di pendenza e segno della curvatura, descrivibili in termini qualitativi quali segmento massimo (un segmento del profilo caratterizzato da pendenza più elevata rispetto ai segmenti successivi e precedenti) o sequenza (una successione di segmenti caratterizzati da una convessità un segmento massimo ed una concavità). Oltre a tale descrizione qualitativa è possibile misurare il valore di pendenza e curvatura dei singoli segmenti.

Questa caratterizzazione del profilo consente di realizzare una classificazione di tipo oggettivo, senza necessariamente dover correlare il tipo di forma individuata alla genesi della stessa (ossia ai processi). Si tratta in effetti di un'analisi parametrica che si discosta dall'analisi interpretativa tipica dello studio di foto aeree o dell'analisi speditiva di carte topografiche ma che porta un importantissimo contributo a quest'ultima; infatti l'informazione che ne scaturisce risulta oggettiva e gode di una maggiore accuratezza descrittiva rispetto al dato oggettivo. In effetti, nell'analisi di tipo soggettivo, la fase di estrazione delle caratteristiche morfometriche e l'interpretazione delle forme costituiscono una fase congiunta. Chi opera delinea le forme di interesse direttamente, seguendo i confini dettati dalla propria esperienza e percezione visiva. Nel caso dell'analisi quantitativa è possibile estrarre i parametri per poi realizzare una carta di sintesi di uno o più parametri aggregati (pendenza, esposizione ecc..) che può facilitare la fase di interpretazione delle forme.

Anche lo studio tramite transetti è spesso usato come strumento di analisi quantitativa. Un transetto è profilo topografico estratto lungo una direzione arbitraria. Lungo tale profilo topografico le misure di angolo di pendenza risultano in pendenze apparenti e non rispecchiano la reale pendenza del rilievo.

Tali tipologie di profilo non sono quindi utilizzabili per l'analisi delle forme. Sono tuttavia utili per trarre diversi tipi di informazioni sulla distribuzione spaziale delle quote. Ad esempio è possibile individuare superfici che sono interrotte da processi di erosione ma topograficamente correlabili come i terrazzi fluviali. Profili trasversali ai fondovalle possono essere utili all'individuazione degli ordini dei terrazzi.

Dipendentemente dallo scopo che si vuole raggiungere tramite l'analisi è quindi necessario scegliere una opportuna direzione lungo la quale estrarre il profilo topografico. La direzione di massima pendenza è una scelta generica che spesso, a causa della limitatezza dei limiti imposti dai metodi manuali di estrazione, non risulta particolarmente efficace per l'analisi delle forme in ampie zone di studio.

In tale contesto particolarmente importante si rivela nell'analisi tradizionale la suddivisione dell'area di studio in unità morfologiche di base. Una scelta utilizzata spesso è quella di analizzare la forma delle singole aree di sottobacino. In tal caso i profili topografici saranno indirizzati lungo le direttrici che meglio possono descrivere la forma del bacino: in un certo senso si tratta di estrarre i profili che racchiudono la maggiore informazione possibile rispetto a tutti i possibili profili estraibili. Risultava infatti troppo oneroso e probabilmente anche di scarsa utilità, estrarre ed analizzare numerosi profili topografici. Nell'epoca del computer è semplice estrarre profili topografici. Chorowitz (1995) ad esempio utilizza l'elaboratore per estrarre in automatico un numero molto elevato di profili. Si tratta di un metodo di analisi differente rispetto ai metodi di analisi bidimensionale della superficie che verranno illustrati nel capitolo 3. La metodologia proposta dall'autore si basa sull'analisi unidimensionale del profilo ed è quindi riconducibile all'analisi tradizionale sopra illustrata ma automatizzata grazie all'uso del GIS. Dalla combinazione dei profili estratti da un DTM viene generata una classificazione dell'intera area di studio in forme secondo sequenze del profilo simili a quelle descritte nel lavoro originale di Young, ad esempio per

rilevare la presenza di elementi particolari come le *cuestas*, caratterizzate da una particolare e caratteristica sequenza di segmenti\*. Poiché il calcolatore può estrarre un numero illimitato di profili dal DTM, è possibile realizzare una classificazione dell'intera area. In pratica senza l'utilizzo dell'elaboratore si rivela necessario una scelta oculata dei transetti. Una scelta suggerita da Young (1972) per un generico sistema vallivo, è quella di considerare punti intermedi tra *talweg* ed *interfluvi*. Il profilo sarà quindi estratto tra i due elementi del bacino passando attraverso il punto seguendo principalmente la direzione di massima pendenza. Suggerisce inoltre di analizzare come dato supplementare i seguenti tipi di profilo estratti lungo determinanti lineamenti del bacino:

- (1) Il profilo longitudinale lungo gli *interfluvi* principali ed altre creste
- (2) Il profilo longitudinale del *talweg*
- (3) La curvatura planare nel punto più acclive di ogni profilo
- (4) L'aspetto del versante nel punto più acclive del profilo ed altri punti se vi è un cambio notevole nella direzione del profilo

E' ovvio che si tratta di indicazioni generali, volte ad una caratterizzazione di una non specifica area di studio, tenendo in considerazione quegli elementi che sono più interessanti dal punto di vista dell'evoluzione di un generico paesaggio vallivo. La scelta del transetto è ovviamente subordinata a l'informazione che si vuole ottenere. Ad esempio Hovius (2000) utilizza i profili dei corsi d'acqua e dei suoi *interfluvi* per analizzare la relazione tra forme fluviali e tettonica, identificando, anche tramite l'ausilio di funzioni interpolanti, l'esistenza di una certa omogeneità nella forma dei profili di numerosi corsi d'acqua presenti nella stessa regione. Profili particolari, detti *swat*, cioè curve che rappresentano il valore medio minimo e massimo dei valori di quota estratti lungo una strisciata longitudinale, sono stati estensivamente utilizzati per l'analisi delle catene

---

\* Tuttavia si evince, analizzando le mappe dei profili estratti in automatico da Chorowitz, che questi sono estratti lungo direzioni che non seguono la direzione di massima pendenza ma piuttosto un angolo di 45 o 90 gradi rispetto alle coordinate del display.

montuose (Fielding 2000, Tippet & Hovius 2000, Ribolini 2000), e la scelta della direzione in questi casi, vista la scala di analisi, è ovviamente lungo direttrici o percorsi che non corrispondono a specifici elementi ma ad esempio lineamenti tettonici, massima estensione della catena ecc..

Tuttavia l'analisi del singolo profilo non è sufficiente per fornire tutte le informazioni sulle caratteristiche morfometriche dell'area di studio, per cui numerosi profili, ed adeguati metodi di analisi di tali rilievi, risultano necessari per caratterizzare la morfologia dell'area. Vari profili devono essere estratti a tal fine. Young raccomanda il rilevamento direttamente sul campo e ciò in un'epoca in cui i calcolatori (e soprattutto i DTM) non erano ancora disponibili come strumento di analisi ne risulta evidente la motivazione: avendo a disposizione come fonte una carta topografica è necessario, per poter scomporre il profilo nelle sue parti caratteristiche risulta necessario effettuare un'interpolazione tra i punti lungo le isoipse intercettate dal profilo topografico. Non avendo a disposizione un computer risultava più speditivo il rilievo topografico di precisione lungo i transetti. In tal modo è possibile inoltre rilevare un gran numero di punti lungo il transetto, molti più rispetto ai punti estraibili dalla carta topografica, risultando quindi possibile l'analisi delle concavità e convessità del profilo in termini quantitativi. A tal fine risulta adeguato l'utilizzo di un'interpolazione lineare, ossia semplicemente congiungere i vari punti con dei segmenti, cosa che si può effettuare manualmente sulla carta millimetrata. Tramite l'analisi dei punti segnati in tal modo è possibile ricavare i valori di pendenza e di curvatura.



## **B.II Modelli digitali della superficie topografica**

### **3.1 Introduzione**

L'insieme dell'archiviazione, la visualizzazione ed analisi della superficie topografica costituisce una delle aree più interessanti tra le svariate applicazioni dei pacchetti GIS. La forma del rilievo analizzata nelle sue componenti quali la, pendenza, l'esposizione e la curvatura, risulta di fondamentale importanza per lo studio di numerosi aspetti dei processi geomorfologici. L'informazione altimetrica della superficie del terreno è gestita all'interno dei software GIS sotto forma di Modello Digitale del Terreno o *Digital Terrain Model* in inglese. Altre sigle utilizzate nell'ambito dell'analisi della superficie terrestre sono DEM o *Digital Elevation Model*, spesso utilizzata come sinonimo della prima ma in effetti riferita ad un tipo di informazione altimetrica differente. Infatti, a seconda del metodo utilizzato per campionare la superficie terrestre, si possono ottenere dati che rispecchiano o meno la reale superficie topografica. Il termine DEM viene spesso utilizzato per indicare un modello della superficie terrestre in presenza di quote rappresentati strutture di tipo antropico o coperture boschive. Da qui la deiezione del termine terreno in favore di elevazione. Un sinonimo per indicare tale tipo di dato è *Digital Surface Model*, utilizzato spesso per i dati derivati da strumentazione laser, ma che in effetti può essere riferita anche a superfici non di tipo terrestre (ad esempio scansioni laser di modelli analogici per la simulazione di processi geomorfici, o nel campo della metrologia [Pike 2000])

### **3.2 Metodi di rappresentazione ed archiviazione dei DTM**

Il modello digitale del terreno è uno esempio di una classe più generale di modelli digitali di fenomeni che si manifestano sotto forma di campi (O'Sullivan & Unwin 2003). Un campo rappresenta un fenomeno che varia in maniera

continua nello spazio. Mentre il modello digitale del terreno rappresenta dati altimetrici, altri modelli digitali possono rappresentare ad esempio valori di temperatura, pressione o pioggia. La quota è appunto un campo di tipo scalare ossia di un fenomeno i cui valori, continui nello spazio, non cambiano se si modifica il sistema di coordinate utilizzato. Un valore scalare descrive solamente l'intensità del fenomeno ad una determinata coordinata spaziale. Un altro esempio di campo scalare è il valore di temperatura dell'aria che consiste in un campo scalare caratterizzato da tre coordinate spaziali. Se consideriamo un volume di aria, ad esempio una regione a forma di cubo, la temperatura all'interno di esso può essere misurata utilizzando uno dei vertici come riferimento per definire il sistema di coordinate spaziali. Né la rotazione né la traslazione del cubo contenente tale volume di aria, operazione che modifica il sistema di coordinate rispetto ad un sistema di riferimento assoluto, influisce sul valore misurato della temperatura alle nuove coordinate ottenute tramite le suddette trasformazioni geometriche. Esistono tuttavia campi scalari il cui valore è ricavato da un campo vettoriale. Ad esempio, la pendenza è un campo scalare, ma esprime la magnitudine di un campo vettoriale, il gradiente della superficie, che essendo un campo vettoriale non risulta invariante alle trasformazioni geometriche. In particolare il gradiente deriva dall'interazione tra un campo scalare, la superficie terrestre, ed un campo vettoriale, il campo gravitazionale, come verrà illustrato nel paragrafo 3.5 parlando dell'analisi differenziale della superficie. Il campo gravitazionale è per semplicità associato all'asse Z (asse verticale) del sistema di coordinate della superficie, per cui ruotando lungo i due assi orizzontali cambia la sia la direzione che il modulo del gradiente. I campi scalari bidimensionali come la quota, vengono gestiti e visualizzati usualmente come superfici. Esistono differenti modi per poter rappresentare ed archiviare una superficie tramite software GIS. Tutti i metodi sono delle rappresentazioni incomplete, ossia non rappresentano in maniera continua il fenomeno spaziale oppure si tratta di approssimazioni, ossia modelli continui ma che si discostano in maniera più o

meno rilevante dai reali valori del fenomeno.

I modi più comuni di rappresentare ed archiviare il dato altimetrico consistono in punti irregolarmente spaziati (o punti sparsi), TIN (Triangulated Irregular Network), linee di livello, punti regolarmente spaziati (*points on regular lattice* sensu Wise 1998) e *raster* (fig. 3.1).

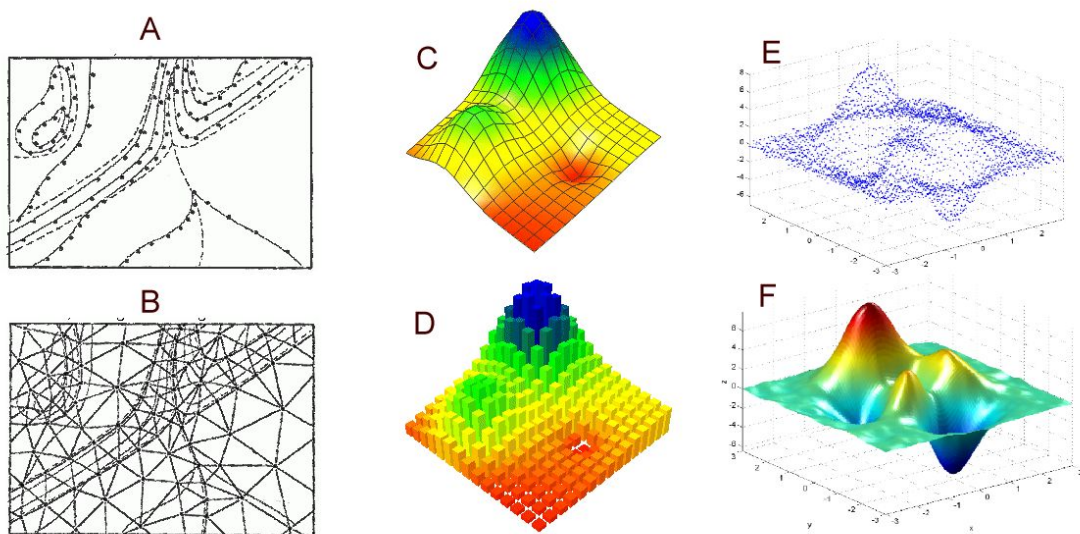


Fig 3.1 Vari modi di rappresentare ed gestire un modello digitale del terreno: (a) Linee di livello (b) Maglia irregolare (TIN) (c) Maglia regolare (punti regolarmente spaziati o regular lattice) (d) raster (e) punti sparsi (f) superficie continua (da Carter 1988, Kidner et alii 1999, FastRBF toolbox manual 2001).

Uno dei più comuni metodi di gestione dei DTM è il modello a celle (di seguito definito come modello *raster*), nel quale l'elevazione è gestita come una struttura formata da celle regolari di forma quadrata che, a meno di approssimazioni dovuti al sistema di proiezione utilizzato, hanno eguale spaziatura verticale ed orizzontale. Un modello analogo ma differente concettualmente è il modello a punti regolarmente spaziati o *gridded* DTM (Wise 1998), che consiste in una serie di punti regolarmente spaziati, e che per semplicità verrà di seguito

denominato modello *grid*. Tali modelli sono diffusi e disponibili a diverse scale. La maggior parte dei dati topografici in formato digitale realizzati dagli istituti nazionali ed internazionali sono disponibili come modello *grid* o *raster*. Tra i modelli disponibili a scala globale vi è il GTOPO30, un modello digitale del terreno realizzato dalla U.S. Geological Survey. Tale modello copre la superficie terrestre da 90 gradi di latitudine nord a 90 sud ed è caratterizzato da una spaziatura della cella di 30 arcosecondi.

Uno dei motivi della grande diffusione dei DTM di tipo *raster* e *grid* (Kidner et alii 1999, Moore et alii 1991, Wise 1998, Wilson & Gallant 2000) è la sua semplice struttura topologica. Difatti tale caratteristica è il motivo della minor diffusione ed applicazione di altri tipi di modelli come ad esempio i TIN. Nonostante siano in grado di generare una rappresentazione del terreno superiore dal punto di vista del dettaglio e della forma delle strutture, questo modello richiede software specifici ed algoritmi di gran lunga più complessi. La TIN fornisce un modello a faccette triangolari. Questi triangoli, che coprono totalmente la superficie dell'area rappresentata, permettono di identificare la quota in ogni punto. Gli altri metodi di archiviazione, *raster* a parte, necessitano di un processo di interpolazione del dato originale (Mitas & Mitasova 1999, Evans 1980, Kidner et alii 1999) per poter identificare i valori intermedi. Il vantaggio principale della TIN è che essendo una struttura di tipo irregolare è rende possibile rappresentare forme che si presentano a scale differenti. Evidenzia inoltre in maniera ottimale le brusche variazioni di pendenza, presenti lungo scarpate o lungo impluvi e displuvi. Questo grazie alla possibilità di utilizzare una differente densità dei nodi, e quindi delle faccette triangolari, lungo la superficie da rappresentare. Per poter racchiudere la stessa quantità di informazione tramite un *raster* o un *grid*, è necessario diminuire la spaziatura rispettivamente delle celle e dei punti fino alla scala di risoluzione della forma che si intende rappresentare, utilizzando quindi un numero di celle molto superiore ai nodi della TIN.

I punti irregolarmente spazati costituiscono un modo molto comune di archiviare il dato spaziale. Per lo più è un modello che deriva da campionamenti iniziali del dato altimetrico. Ad esempio può costituire un dato grezzo acquisito tramite un sistema di scansione laser; non può essere considerato un vero modello della superficie perché esso non fornisce alcuna informazione sulla natura della superficie tra i punti dati. Gli altri modelli citati possono essere derivati direttamente - tramite interpolazione - da punti irregolarmente spazati, tranne il modello a linee di livello. Nel caso delle linee di livello il dato altimetrico risulta costante lungo una curva. Tale struttura può essere gestita in vari formati vettoriali, come ad esempio polilinee 2d o 3d (ArcView 1993). Anche in questo caso il dato altimetrico non è restituibile ovunque e quindi la superficie terrestre non risulta definita ovunque.

Per analizzare la superficie topografica risulta quindi necessario utilizzare un metodo di interpolazione (ad esempio a partire da linee di livello) che restituisca un modello di superficie continua da cui è possibile estrarre i valori di quota in ogni punto. I modelli di superficie continua non sono utilizzati normalmente per archiviare e visualizzare il DTM. Come vedremo, il modello a superficie continua è un modello di tipo derivabile ed è utilizzato al fine della analisi della superficie.

Il modello di tipo *grid* è costituito da un'insieme di punti regolarmente spazati che derivano dalla valutazione lungo una griglia dei valori derivanti da interpolazione di punti irregolarmente spazati o da curve di livello. Tali punti sono interpretati come una maglia regolare analoga alla struttura dei TIN e sono un modello "lattice" cioè una maglia poligonale costituita da nodi e segmenti. Rispetto alle TIN, tale struttura è caratterizzata da nodi regolarmente spazati, e di conseguenza faccette costituite da quattro vertici. Questo tipo di modello è di più semplice gestione rispetto alle TIN. La semplicità deriva dal fatto che tale struttura può essere convertita in un *array* bidimensionale e quindi in una struttura di tipo *raster*. Gli *array* bidimensionali o matrici, sono facilmente manipolabili tramite un qualsiasi linguaggio di script o di programmazione, quale

il Fortran (Comincioli 1991). Di conseguenza la sintassi necessaria a manipolare il dato altimetrico somiglia alla sintassi dell'algebra matriciale per le moltiplicazioni e le addizioni. Quasi tutti i GIS che contemplano la gestione del dato raster, sono forniti di funzioni che consentono le operazioni di *map-algebra* (cfr. paragrafo 4.5). La *map-algebra* (Shapiro & Westerwelt, 1992) costituisce un vero e proprio linguaggio di script che permette di manipolare qualsiasi dato raster tramite operatori logici e matematici.

Di contro gli algoritmi necessari a gestire linee di livello e TIN sono più complessi poiché è più complessa la struttura del dato. Ad esempio un TIN non può essere convertito in *raster* in maniera immediata ma necessita di un processo di “campionamento”; poiché la superficie delle faccette triangolari è continua è necessario ricavare una quota da assegnare ad ogni cella del *raster*, generalmente il valore alle coordinate del centro della cella. Data la maggiore complessità necessaria a manipolare gli altri tipi di modelli menzionati, sono stati sviluppati per i DTM *raster* e *grid* un numero molto maggiore di algoritmi. Pertanto tali formati sono sfruttati in maniera più estesa rispetto agli altri.

Un'altra importantissima ragione della popolarità di tali modelli è la compatibilità con il formato di rappresentazione ed archiviazione di altre tipologie di dati spaziali. In particolare, il DTM *raster* ha una struttura dati analoga a qualsiasi altro strato informativo derivato da un'insieme di valori di un fenomeno spaziale rappresentato da celle regolarmente spaziate. Alla semplicità e compatibilità dei DTM *raster* si contrappone una certa ambiguità nell'interpretazione del dato. Nei DTM *raster*, difatti, è definita solamente la struttura dell'insieme dei dati, struttura che non definisce in maniera implicita la localizzazione del dato altimetrico come nel caso del modello *grid*.

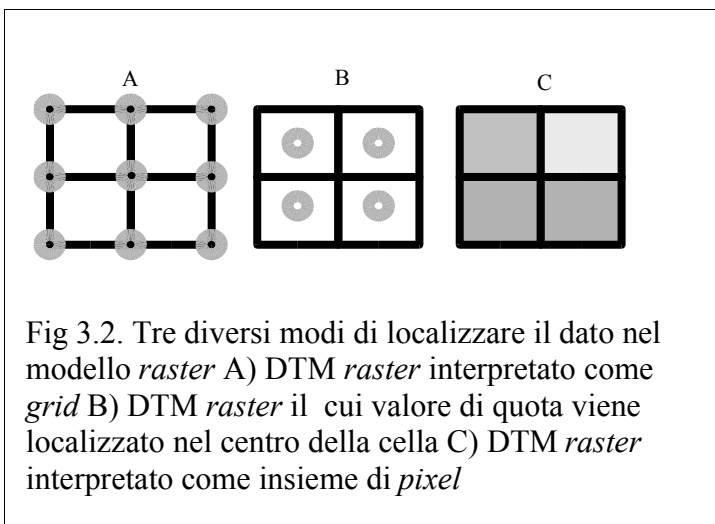
Nel modello *grid* i valori sono gestiti come punti geometrici definiti tramite le loro coordinate geografiche. Spesso i due termini *raster* e *grid* vengono usati indistintamente per identificare lo stesso tipo di modello DTM. Tuttavia più in generale si conviene che la cella del *raster* non rappresenta un punto geometrico,

ma un'area attraversata da una certa quantità di informazione. In un'immagine digitale, ad esempio, tale informazione è l'intensità luminosa. Nel caso del DTM è il valore altimetrico riferito all'intera area di una cella. Tale modo di interpretare il dato, analogo appunto al sistema di archiviazione delle immagini digitali, è quello più propriamente definito come DTM *raster* (o *grid* in *pixel mode* secondo Wise 1998). Questa distinzione risulta importante quando si affronta ad esempio il problema del ricampionamento (cioè del cambiamento di risoluzione) del dato e della interpolazione ed della conversione tra i due formati. Spesso il DTM è il risultato di un'operazione di interpolazione, per cui i valori associati alle celle non sono realmente valori derivati da punti campionati sulla superficie reale. Solo i punti dati (cioè i punti lungo le linee di livello, punti derivati da scansioni LIDAR ecc..) che vengono utilizzati dall'algoritmo di interpolazione per generare il *raster* od il *grid*) sono punti realmente campionati mentre i valori dei punti del *grid* e le celle del *raster* del DTM interpolato non derivano da valori realmente misurati sul campo ma, “valutati” lungo intervalli regolari. Dall'interpolazione si ottiene invece una superficie continua. Mentre nel caso del *grid*, il valore si riferisce ad un punto geometrico della superficie interpolata, nel caso del *raster*, a seconda del software utilizzato, alla intera area della cella viene attribuito il il valore della superficie valutato alle coordinate spaziale del centro geometrico della stessa. Quest'ultima opzione è quella utilizzata più di frequente data la semplicità di implementazione (infatti nel primo caso sarebbe necessario calcolare l'integrale della funzione interpolante) per cui i valori valutati coincidono in entrambi i formati, ma, questo tipo di operazioni non sono trasparenti all'utente del GIS, che dovrebbe quindi sempre informarsi (ma ciò non è sempre possibile, specie per i software commerciali) sul funzionamento dell'algoritmo di interpolazione. E' possibile infatti che il software utilizzato utilizzi il valore valutato alle coordinate di un angolo della cella. Nel caso di DTM di cui non si conosce bene il metodo con cui sono stati realizzati, soprattutto per i modelli reperiti ad esempio via internet è necessario quindi

tenere conto di alcuni aspetti critici. Se un modello *raster* viene interpretato come *grid*, l'ubicazione dei valori risulta comunque sconosciuta. Come detto in precedenza il *grid* non è costituito da un'insieme di valori associati ad aree, ma da punti distanziati in maniera regolare. Nonostante si tratti di un'insieme di punti, risulta più efficiente archiviare tale struttura sotto forma di matrice piuttosto che come una serie di punti definiti esplicitamente. Un formato di archiviazione esplicito è chiamato comunemente x,y,z ed è costituito da un *file ascii* ove per ogni riga di testo vengono registrate le tre coordinate spaziali. Questo tipo di archiviazione del DTM suggerisce che il dato debba essere gestito come formato vettoriale piuttosto che come formato *raster*. Risulta molto più compatto tuttavia archiviare tale DTM sfruttando la proprietà di spaziatura regolare tra i punti. Un *file ascii* di questo tipo è , ad esempio, quello utilizzato dalla ESRI (ArcView 1993). In questo formato, nel *file* sono registrate, in un cosiddetto *header*, le coordinate geografiche dei vertici della regione rettangolare ricoperta dal *raster*, il valore di spaziatura ed il numero di righe e colonne di tale regione. In tal modo non è più necessario registrare le due coordinate geografiche della singola cella ma solo il suo valore realizzando un *file* più compatto.

Nel modello *raster*, il dato altimetrico è associato all'area dell'intera cella ma per convenienza la sua ubicazione è identificata spesso come il centro del pixel (indipendentemente da come sia stato valutato il dato interpolato). Questo espediente rende semplice valutare ad esempio le distanze tra tematismi, come strade ed edifici, quando questi sono archiviati in formato *raster*, anche se è generalmente riconosciuto che tali distanze hanno un elemento di imprecisione dovuto all'incertezza dell'ubicazione geografica del dato.





Il problema é che realmente pochissimi software GIS contemplano a livello logico qualsiasi distinzione tra un modello *grid* ed i modelli *raster*, ma solo tra dato *raster* e dato vettoriale, formato che include la gestione dei punti.

Trattare i modelli *raster* e *grid* come se fossero rappresentazioni equivalenti dà vita quindi a problemi di ambiguità sia sulla localizzazione del dato che sulla sua estensione areale (fig 3.2).

Questi distinguo sono di importanza notevole poiché la maggior parte dei GIS è dotata di funzioni che consentono di convertire un dato *raster* in un'insieme di punti regolarmente spazati e viceversa. Inoltre alcune funzionalità, quali il cambiamento di scala del DTM o alcuni algoritmi di interpolazione, sono disponibili solo per un tipo di modello e non per l'altro. Risulta quindi importante conoscere come il dato viene gestito dal software utilizzato.

### 3.3 Cenno sui metodi di acquisizione

Il primo passo per la generazione di un DTM è ovviamente il campionamento dei punti a terra. Esistono diversi mezzi per campionare le quote della superficie terrestre. Il metodo più diretto consiste nell'utilizzare una stazione totale ossia un

sistema computerizzato in grado di acquisire dati da teodoliti e distanziometri. Il sistema della triangolazione consente di procedere secondo una maglia a partire da una base ossia una coppia di punti dalle coordinate note. Si tratta ovviamente di un metodo non molto speditivo ma molto accurato se integrato da punti di appoggio o sistemi GPS di tipo differenziale. Uno dei limiti maggiori è quello dell'accessibilità del territorio. I metodi di telerilevamento sono metodi che non prevedono il rilievo diretto a terra ma prevedono un sensore montato su un apparecchiatura mobile aerea o spaziale.

### ***Aerofotogrammetria***

Si tratta di un metodo utilizzato da anni per l'acquisizione rapida di quote. Si basa sul principio del parallasse (Perlant 2000). Un sensore montato su di un aereo, in genere una macchina fotografica con pellicola ad alta sensibilità nel campo del visibile o del vicino infrarosso, viene utilizzata per scattare fotogrammi secondo un piano di volo che consente di sovrapporre almeno il 60% di due fotogrammi adiacenti. Ogni coppia di fotogrammi costituisce una stereocoppia che consente di ottenere immagini prospettiche. Per la generazione delle quote a partire da stereocoppie è possibile utilizzare sia sistemi analitici che digitali. Nei sistemi analitici l'operatore utilizza un sistema stereoscopico dotato di pennino luminoso che sposta lungo tre assi virtuali fino a far coincidere il punto con quello che egli percepisce visivamente come la superficie del terreno. Nei sistemi digitali si impiega un software specifico che utilizza il principio della autocorrelazione delle immagini per individuare automaticamente gli stessi punti sui due fotogrammi. Entrambi i metodi sono in grado di generare rapidamente un gran numero di punti ma sono soggetti ad alcune limitazioni come ad esempio la presenza di vegetazione o strutture antropiche. L'errore in quota è variabile a seconda dell'andamento del rilievo, dalla presenza di ombre nei fotogrammi e dalla presenza o meno di tessitura.

## ***Radar***

Lo sviluppo dei sistemi satellitari ha permesso di sfruttare il metodo dell'interferometria radar, già utilizzato in passato per ottenere immagini tridimensionali della superficie lunare, per la produzione di DTM della superficie terrestre. Il vantaggio rispetto ai sistemi che si basano su radiazioni nel campo del visibile o del vicino infrarosso, è quello di essere esente da disturbi dovuti alle differenti condizioni atmosferiche e dalla presenza o meno della luce solare. Il radar è costituito da un emettitore di onde nel campo delle microonde, che invia un fascio di impulsi verso terra, e da un sistema in grado di rilevare la radiazione riflessa dal terreno. Misurando il tempo di ritorno è possibile ricavare la distanza. La tecnica usata è quella del SAR o *Syntetic Aperture Radar* (Ferri 1996), che può essere montato sia su satellite che su aeromobile e sfrutta il moto della piattaforma per simulare la presenza di un'antenna molto più lunga ed in grado di ricevere più impulsi con angolazioni differenti dello stesso punto del terreno. Il principio su cui si basa la misurazione della quota è quello dell'interferometria radar che utilizza lo stesso principio geometrico della aerofotogrammetria, cioè osservare due punti della superficie da posizioni leggermente sfalsate. Benché il principio geometrico sia simile, l'interferometria radar differisce nel principio fisico che si basa sulla differenza di fase tra due impulsi.

Nel febbraio del 2000 una missione dello *Space Shuttle della National Aeronautics and Space Administration* (NASA), in collaborazione con la *National Imagery and Mapping Agency* (NIMA), l'agenzia spaziale tedesca (DLR) e l'agenzia spaziale italiana (ASI), ha scansionato l'80% della superficie terrestre e realizzato un DEM tramite la tecnica della interferometria radar. Il nome della missione SRTM (*Shuttle Radar Topography Mission*) ha generato una mole di dati enorme (10 terabyte) che sono stati quindi elaborati e resi pubblici in due formati differenti SRTM-1 (circa 30 m di spaziatura orizzontale) per il territorio degli USA e SRTM-3 (circa 90 m di spaziatura) per il resto della superficie terrestre.

## ***Laser***

L'*Airborne Laser Scanning* è un sistema composto da uno scanner laser denominato LIDAR *Light Detectionon And Ranging* (Naesset 1997) montato su di un aeromobile. Il sistema integra componenti al suolo con il modulo in volo. Un impulso laser viene inviato al suolo mantenendo l'aereomobile perpendicolare al suolo grazie ad un sistema di navigazione inerziale, ed il segnale riflesso viene registrato. La distanza tra lo scanner montato sull'aereomobile ed il suolo viene misurata calcolando con grande precisione (tramite orologio atomico) il tempo di ritorno dell'impulso. Al suolo vi è una stazione DGPS che consente di registrare le coordinate del punto misurato. E' possibile ottenere misure della quota con precisione di  $\pm 15$  cm. Tale sistema ha il vantaggio di non essere affetto dal problema della presenza di ombre mentre ha lo svantaggio di misurare la quota di alberi e strutture il che impedisce la misurazione della reale superficie topografica anche se recentemente laser dalla potenza più elevata (laser di classe IV) sono in grado di attraversare il denso fogliame e di utilizzare i diversi impulsi di ritorno, cioè quelli provenienti dalla chioma e dal suolo, per misurare sia la quota della copertura arborea che la reale superficie topografica. Oltre che misurare la quota, il sensore è in grado di produrre un'immagine fotografica utilizzando il valore di riflettanza dell'impulso di ritorno consentendo di realizzare di fatto delle vere e proprie ortofoto.

## ***Acquisizione da carte topografiche***

Il metodo più diffuso ed economico per la creazione dei DTM consiste nell'utilizzare le carte topografiche disponibili. Il *dataset* GTOPO30 (EROS data center, USGS, <http://edcdaac.usgs.gov/gtopo30>), che ricopre buona parte della superficie terrestre (da 90 gradi sud a 90 nord) con una risoluzione orizzontale di 30 arcosecondi (circa un 1km di spaziatura orizzontale), o il DTED, disponibile a vari livelli di risoluzione, sono stati realizzati acquisendo le linee di livello di carte topografiche di varia provenienza e fondendo i DTM derivati dalle singole

mappe, correggendo eventuali problemi dovuti ai diversi sistemi di proiezione. Nel caso del GTOPO30, ad esempio, il sistema di coordinate è un longitudine-latitudine proiettato sullo sferoide WGS84. L'errore sia planimetrico che di quota varia a seconda della latitudine e della fonte cartografica disponibile per le diverse aree del pianeta. Nel caso delle tavolette USGS del territorio degli USA il datum è UTM-WGS84. Le caratteristiche di accuratezza verticale sono disponibili sui siti internet delle rispettive agenzie, dove è possibile scaricare gratuitamente tali DTM.

Denominazione dataset	Spaziatura orizzontale	Copertura
ETOPO5	5 arcominuti	Globale incluse le batimetrie
<a href="#">TERRAIN BASE</a> (NGDC)	5 arcominuti	Globale incluse le batimetrie
<a href="#">GTOPO30</a> (USGS/EROS DATA CENTER)	30 arcosecondi	Globale
<a href="#">DTED LEVEL 0</a> (NIMA)	30 arcosecondi	50% superficie terrestre
<a href="#">GLOBE</a> (NGDC)	30 arcosecondi	Globale incluse le batimetrie
<a href="#">1:250,000 DEM</a> (USGS disponibili da vecchi dati NIMA)	3 arcosecondi	Intero territorio USA
<a href="#">1:24,000 DEM</a> (USGS)	30 m proiezione UTM	Molte aree USA
<a href="#">1:24,000 DEM</a> (USGS)	10 m proiezione UTM	Alcune aree USA
SRTM-1 (NASA/NIMA)	30 m proiezione UTM	Intero territorio USA
SRTM-3 (NASA/NIMA)	90 m proiezione UTM	80% della superficie terrestre

*Tab 3.1 – I più diffusi dataset disponibili gratuitamente tramite World Wide Web (da Guth 1999, modificato)*

Il metodo per acquisire il dato topografico si basa sulla digitalizzazione delle linee di livello. Ogni linea di livello viene convertita in una polilinea. Una polilinea è dato vettoriale costituito da nodi che congiungono segmenti. Queste polilinee possono essere generate tramite un tavolo digitalizzatore o tramite un apposito software che lavora su una immagine scansionata della carta topografica. In ambiente GIS è possibile lavorare utilizzando tali immagini scansionate e generare le polilinee in maniera manuale, cioè seguendo con un puntatore le curve di livello, o in maniera semiautomatica, tramite appositi algoritmi in grado di inseguire le linee. Alla polilinea generata viene quindi

assegnato il suo valore di quota in un apposito campo del database associato al dato vettoriale. Questo campo verrà utilizzato per l'interpolazione al fine di generare il modello digitale del terreno.

Ogni metodo consente di ottenere un certo numero di campioni secondo uno schema voluto o, molto più frequentemente, derivante dal metodo utilizzato. I campionamenti più diffusi sono di tipo regolare, casuale e casuale stratificato. Il campionamento regolare garantisce una buona distribuzione sopra l'area anche se potrebbe essere soggetto ad errori dovuti a fluttuazioni periodiche della superficie. E' preferibile spesso avere a che fare con qualche tipo di campionamento casuale. Quello di tipo puro porta spesso all'agglomerazione di punti in *cluster* per cui si preferisce utilizzare il campionamento casuale stratificato. Ovviamente scegliere il tipo di campionamento non è sempre possibile. Ad esempio il campionamento da satellite o da LIDAR consentono di ottenere una fitta maglia più o meno regolare. Se la fonte a disposizione per la produzione di DTM sono carte topografiche, il campionamento non è né random né regolare ma segue l'andamento delle curve di livello.

### **3.4 Metodi di interpolazione per la generazione di DTM**

L'interpolazione spaziale consiste in un metodo per ottenere valori di punti intermedi di un determinato fenomeno geografico. Data la sua importanza il concetto di interpolazione spaziale è un argomento affrontato in diversi testi specialmente rivolti all'aspetto numerico del problema (es. Mitsova & Mitsova 1993, Comincioli 1995). La formulazione del problema è la seguente:

dati  $N$  valori di un fenomeno spaziale

$$Z_j, \quad j = 1, \dots, N$$

misurati nei punti discreti

$$\mathbf{r}_j = (x_j^{(1)}, x_j^{(2)}, \dots, x_j^{(d)}), \quad j = 1, \dots, N$$

all'interno di una determinata regione di uno spazio  $d$ -dimensionale, trovare una funzione  $d$ -variata  $F(\mathbf{r})$  che passi tra i punti dati, cioè soddisfi la condizione

$$F(\mathbf{r}_j) = z_j \quad j = 1, \dots, N$$

Le soluzioni anche per un numero di punti molto basso sono infinite a meno di non introdurre delle condizioni ulteriori. Si determina quindi la necessità di scegliere il metodo di interpolazione. I metodi di interpolazione possono essere classificati in base all'approccio utilizzato (geometrico, matematico, statistico, fisico ecc..) o al numero di campioni che utilizza per risolvere la superficie.

Quest'ultimo criterio di classificazione porta a distinguere i metodi di interpolazione globali, cioè che utilizzano i valori di tutti i punti campione, da quelli di tipo locale, in cui si utilizza solamente un certo numero di punti vicini per calcolare il valore di un determinato punto. Quasi tutti i metodi di interpolazione globale tuttavia possono utilizzare entrambi gli approcci grazie al fatto che una superficie che descrive un fenomeno reale risulta spesso dotata di proprietà geometriche locali. In particolare per fenomeni spaziali dotati di forte autocorrelazione, quali la distribuzione delle quote di una superficie topografica, è possibile risolvere la superficie interpolante per pezzi.

### 3.4.1 Metodi globali

L'approccio globale consiste nell'utilizzare tutti i punti campionati per ottenere la superficie interpolante. Si definisce quindi una funzione della quota, come ad esempio una polinomiale del tipo

$$F(z) = a + b x + c y + d xy + e x^2 + f y^2$$

dove  $a, b, c, d, e, f$  sono i coefficienti del polinomio, che risulta impiegando i valori di tutti punti campione, consente di ottenere una superficie che si adatta ai punti dati. Il problema è che tale equazione possiede una soluzione esatta solo se i punti dati sono pari al numero di coefficienti del polinomio. Risolvendo il relativo sistema di equazioni è possibile ottenere i valori dei coefficienti e successivamente, sostituendo all'equazione così ottenuta i valori di  $x$  e  $y$  di un qualsiasi punto, è possibile ottenere i valori incogniti di  $Z$ . Se si utilizzano un numero di punti superiore al numero dei coefficienti del polinomio, le soluzioni diventano infinite ed è quindi necessario imporre un ulteriore criterio, ovvero scegliere il metodo di interpolazione. Nel caso della polinomiale, per ottenere una funzione interpolante che passi per tutti i punti campione bisogna utilizzare un polinomio di grado pari al numero di punti. Risulta evidente come tale metodo non sia utilizzabile se il numero di punti è elevato poiché diventa difficile se non impossibile risolvere il relativo sistema di equazioni. Si può ricorrere così a sistemi che siano risolvibili a scapito dell'esattezza, in cui la funzione interpolante non passi esattamente per i punti dati. Il metodo dei minimi quadrati consente di ottenere una superficie, detta superficie di regressione, utilizzando i valori dei punti campione ma individuando una funzione che approssima il loro andamento. Il caso più semplice è quello della regressione lineare. Nel caso bidimensionale la formula risulta la seguente

$$f(x,y) = a + bx + cy$$



dove  $f(x)$  è la funzione di regressione (multipla in questo caso) che individua un piano definito dai coefficienti  $a$ ,  $b$  e  $c$ .

E' possibile ottenere una migliore approssimazione utilizzando polinomi di ordine superiore. Ad esempio un polinomio di secondo ordine quale il seguente

$$f(x,y)=b_0+b_{01}x+b_{10}y+b_{11}xy+b_{20}x^2+b_{02}y^2$$

$$Z_i=b_0+b_{01}x+b_{10}y+b_{11}xy+b_{20}x^2+b_{02}y^2+e_i$$

dove  $f(x,y)$  è la superficie stimata, mentre  $Z_i$  rappresenta il valore reale della quota nel punto dato.

Poichè si tratta di una funzione in  $x,y$  si otterranno i singoli valori di quota  $Z_i$  ed i relativi residui  $\epsilon_i$  (cioè di quanto si discosta la superficie quadratica dai punti utilizzati per l'interpolazione). In sintesi:

$$Z_i = f(x_i, y_i) + \epsilon_i$$

Che è esprimibile in forma matriciale come

$$\mathbf{z} = \mathbf{X} \mathbf{b} + \boldsymbol{\epsilon}$$

Con  $\mathbf{X}$  matrice delle variabili indipendenti:

$$\mathbf{X} = \begin{pmatrix} 1 & x_1 & y_1 & x_1 y_1 & x_1^2 & y_1^2 & \dots \\ 1 & x_2 & y_2 & x_2 y_2 & x_2^2 & y_2^2 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & x_n & y_n & x_n y_n & x_n^2 & y_n^2 & \dots \end{pmatrix}$$

$\mathbf{b}$  vettore dei coefficienti  $[b_0, b_{01}, b_{10}, \dots]$  ed  $\boldsymbol{\epsilon} = [\epsilon_1, \epsilon_2, \dots, \epsilon_n]$  vettore dei residui.

Per miglior grado di approssimazione si intende il ridurre la somma dei quadrati

dei residui al minimo possibile. E' dimostrabile che tale minimo si ottiene se è soddisfatta la seguente condizione di ortogonalità (O'Sullivan & Unwin 2003, Comincioli 1995)

$$(\mathbf{X}\mathbf{b})^T(\mathbf{y}-\mathbf{X}\mathbf{b})=0$$

Dove  $\mathbf{y}$  è il vettore delle variabili dipendenti (gli  $z_i$ ) per cui i coefficienti possono ricavarsi risolvendo il sistema

$$\mathbf{b}=(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}$$

che porta alla soluzione ai minimi quadrati in quanto tale soluzione minimizza la somma dei quadrati dei residui.

Tale tecnica di interpolazione viene utilizzata spesso per ottenere il *trend* (da cui il nome *Trend Surface*) del fenomeno oggetto di studio, ossia l'andamento globale. All'aumentare del grado del polinomio la superficie tende ad abbandonare il *trend* poiché con un numero via via maggiore di coefficienti approssima sempre meglio i punti campione e tende ad evidenziare maggiormente le variazioni locali della superficie piuttosto che l'andamento globale. Ciò, in combinazione con la proprietà di autocorrelazione della superficie topografica, permette di risolvere il problema dell'interpolazione per pezzi. Infatti la quota non è una funzione casuale ma piuttosto i punti vicini risultano più correlati di punti lontani (Isaaks & Srivastava 1989, O'Callagan & Unwin 2002). Diminuendo il numero di punti utilizzati per l'interpolazione, è sufficiente un polinomio di grado inferiore per ottenere lo stesso livello di approssimazione ottenibile con un polinomio di grado elevato ma che utilizza tutti i punti della regione e di conseguenza una minore complessità di calcolo. Molte altre funzioni diverse dalla polinomiale sono caratterizzate da un comportamento locale. Per sfruttare il metodo di interpolazione per pezzi è necessario utilizzare opportune tecniche di segmentazione in modo che la superficie interpolata risulti continua, ossia che i segmenti non siano visibili ma

connessi a formare un'unica superficie continua (Cebecauer et alii 2002)

### ***Metodo variazionale***

L'approccio variazionale descritto da all'interpolazione e all'approssimazione si basa sull'assunto che la funzione interpolante debba passare attraverso (o molto vicino) ai punti dati e, allo stesso tempo, debba essere il più possibile regolare (la cosiddetta “smoothness”). Il termine variazionale si riferisce al fatto che modificando i vincoli ed i parametri della funzione si ottiene una variazione di energia della funzione interpolante, energia che viene considerato indice di regolarità della superficie. Questi due requisiti sono combinati in una unica condizione di minimizzazione sia delle somme delle deviazioni dai punti misurati, che permette di variare il grado di approssimazione dell'interpolante, che della seminorma della funzione, che definisce la parte variazionale. L'energia+la deviazione dai vincoli (punti dati) deve essere minimizzata (Mitas & Mitsova 1993):

$$\sum_{j=1}^N |z_j - F(r_j)|^2 w_j + w_o I(F) = \min$$

dove  $w_j$  e  $w_o$  sono dei pesi positivi e  $I(F)$  denota la seminorma della funzione che consiste in un integrale di una funzione (Tab 3.2).

metodo	$I(F)$	Forma di Eulero-Lagrange
Membrana	$\int [F_x^2 + F_y^2] d\mathbf{r}$	Armonica
Min. Curvatura, thin plate spline (a)	$\int [F_{xx}^2 + F_{yy}^2 + 2F_{xy}^2] d\mathbf{r}$	Biarmonica
Thin plate spline+tension (b,c)	$\int \varphi^2 [F_x^2 + F_y^2] + [F_{xx}^2 + \dots] d\mathbf{r}$	Armonica + biarmonica
Thin plate spline regolarizzate (c)	$\int [F_{xx}^2 + \dots] + \tau^2 [F_{xxx}^2 + \dots] d\mathbf{r}$	Biarmonica +6° ordine
TPS regolarizzate +tensione (d)	$\sum_{mn} c_{mn}(\varphi) \int [F_{x^m y^n}]^2 d\mathbf{r}$	Ogni ordine

<sup>a)</sup> (Briggs 1974, Duchon 1975, Wahba 1990, Hutchinson and Bischof 1983)

<sup>b)</sup> (Franke 1985, Hutchinson 1989)

<sup>c)</sup> (Mitas and Mitasova 1988)

<sup>d)</sup> (Mitasova et al. 1995, Mitas and Mitasova 1996)

Tab. 3.2 Modificato da Mitas & Mitasova 1999

La soluzione di tale equazione, sotto condizioni generali, può essere espressa come una somma di due componenti (Talmi & Gilat 1977):

$$F(\mathbf{r}) = T(\mathbf{r}) + \sum_{j=1}^N \lambda_j R(\mathbf{r}, \mathbf{r}_j) \quad (3.1)$$

dove  $T(\mathbf{r})$  è la "trend function", generalmente un polinomio di basso grado, e  $R(\mathbf{r}, \mathbf{r}_j)$  è una funzione a base radiale con associati dei pesi  $\lambda$ , ed ha una forma dipendente dalla scelta di  $I(F)$ .

Una seminorma bivariata con quadrati delle derivate seconde determina una funzione *thin plate spline* che minimizza la curvatura della superficie e imita una lamiera di acciaio costretta a passare attraverso determinati punti: la forma di equilibrio della lamiera minimizza l'energia di deformazione che è strettamente correlata alla curvatura della superficie. A causa della rigidità della superficie si ottengono sovrastime e sottostime dei valori (cosiddetti *overshoot* e *undershoot*) in aree dove i dati determinano bassi gradienti e le derivate di secondo ordine divergono nei punti, rendendo difficile l'analisi. Il problema dell'*overshoot* ed *undershoot* può essere attenuato integrando le derivate di ordine primo nella seminorma  $I(F)$  determinando così una funzione *thin plate spline* con tensione. Le proprietà analitiche (in particolare la derivabilità) di questa funzione vengono

migliorate aggiungendo le derivate di ordine superiore nella seminorma  $I(F)$ , portando così ad una funzione con derivate seconde, terze ecc..., ossia una funzione derivabile infinitamente e ciò rende la funzione regolare in tutto l'intervallo di definizione (*Regularized Spline with Tension*, o RST di Mitas & Mitasova). A seconda della seminorma  $I(F)$  la soluzione è determinata da una funzione a base radiale differente.

Una funzione a base radiale (*Radial Basis Function* o RBF) è un metodo utilizzato in molti campi, tra cui la soluzione di problemi di interpolazione, la regressione non lineare e la costruzione di classificatori non lineari come ad esempio le reti neurali artificiali (Webb 2002). La formula (3.1) può essere riscritta utilizzando la notazione di (Carr et alii 2002):

$$F(\mathbf{x}) = p(\mathbf{x}) + \sum_{i=1}^N \lambda_i R(|\mathbf{x} - \mathbf{x}_i|), \quad \mathbf{x} \in \mathbb{R}^d$$

dove  $p(x)$  è un polinomio di grado almeno  $k$ ,  $\lambda_i$  è un valore reale pesato e  $R(\mathbf{x})$ , posto  $\mathbf{x} = |\mathbf{x} - \mathbf{x}_i|$ , è una funzione base di un valore reale positivo  $\mathbf{x}$ , dove  $\mathbf{x}_i$  è un vincolo o “centro” della funzione. Le funzioni base più utilizzate sono la “thin plate spline”,  $R(x) = x^2 \log(x)$ , la triarmonica  $R(x) = x^4 \log(x)$ , la gaussiana,  $R(x) = \exp(-cx^2)$  e la multiquadratica,  $R(x) = \sqrt{(x^2 + c^2)}$ .

Per le funzioni di tre variabili si utilizzano la biarmonica  $R(x) = x$ , e la triarmonica  $R(x) = x^2$ . Nel caso delle *Regularized Thin Plate Spline* con tensione (Mitasova & Mitas 1993) la soluzione consiste nel porre  $p(x) = a_1$  ovvero un valore costante del “trend” ed la funzione a base radiale assume la seguente forma

$$R(\mathbf{x}_j) = - \left\{ \ln \left[ \left( \frac{\varphi \mathbf{x}_j}{2} \right) \right] + E_1 \left[ \left( \frac{\varphi \mathbf{x}_j}{2} \right)^2 \right] + C_E \right\}$$

dove  $\varphi$  è un parametro di tensione  $C_E$  è la costante di Eulero ed  $E_1$  è l'integrale della funzione esponenziale.

L'algoritmo presente nel gis GRASS è caratterizzato da differenti parametri che consentono di modificare il risultato finale dell'interpolazione.

Modificando il coefficiente di tensione si modifica il comportamento della superficie da una membrana elastica ad una lamiera rigida. L'incremento di tale parametro ha l'effetto di ridurre l'influenza dei punti distanti così che il valore del punto stimato sarà maggiormente influenzato dai punti vicini. Al contrario una tensione bassa ha l'effetto di aumentare il raggio di influenza e quindi l'attrazione dei punti distanti. Questo parametro agisce quindi come fattore di scala della funzione. L'interpretazione dal punto di vista fisico è più intuitiva poiché la tensione modifica il comportamento della superficie costretta a passare tra i punti dati. Una tensione elevata porta alla realizzazione di una superficie simile ad una membrana elastica tirata tra vincoli puntiformi (i punti dati), e la sua visualizzazione evidenzia crateri e cuspidi. Una tensione bassa equivale ad una lamiera metallica rigida che si flette tra i punti dati generando a volte dei problemi se i punti adiacenti sono caratterizzati da rapidi cambi di quota. Tale problema è discusso nel paragrafo 3.7. Lo *smoothing* consente di aggiungere una tolleranza rispetto alla precisione con cui la superficie attraversa i punti dati. In tal modo è possibile, aumentando tale parametro, modificare il comportamento della funzione da interpolante esatto, a *trend surface*.

Le RBF sono funzioni utilizzate per la soluzione di problemi di interpolazione di punti sparsi poiché il sistema di equazioni lineari associato è invertibile sotto ogni condizione di localizzazione dei punti dati.

La soluzione risulta però essere costosa in termini di tempi di calcolo e risulta quindi essenziale l'individuazione di un algoritmo in grado di eseguire in tempi accettabili tali calcoli. L'algoritmo presente in GRASS (s.surf.rst), utilizza una tecnica di segmentazione che permette di suddividere i punti secondo aree che si sovrappongono, in modo da ottenere un collegamento regolare tra le porzioni di

superfici adiacenti. Il metodo della segmentazione comporta una criticità nella scelta dei parametri. Nel caso dell'algoritmo s.surf.rst, i parametri consistono nel numero di punti inclusi in ogni segmento (segmax), e nel numero di punti totali (npmin) utilizzati per interpolare i pezzi di superficie. Una tassellazione della superficie risulterà evidente in casi dove i punti sono ad esempio molto sparsi o i parametri di tensione e smoothing non sono utilizzati correttamente. Beatson et alii (1999) hanno invece sviluppato un metodo che consente di ridurre la complessità computazionale della soluzione del sistema di equazioni riducendolo da  $O(N^2)$  a  $O(N \log_2 N)$ . In questo modo è possibile risolvere per intero, cioè senza utilizzare la tecnica della segmentazione, la superficie interpolante. Tale metodo è stato implementato in un tool commerciale (FastRBF 2003) in grado di risolvere, per le superfici, le funzioni Thin Plate Spline, la triarmonica e la multiquadratica.

### **3.4.2 Metodi locali**

Tali metodi si basano sull'assunto che ogni punto influenza la superficie risultante solo fino ad una certa distanza finita. I valori in punti non campionati differenti vengono calcolati tramite funzioni con parametri differenti e le condizioni di continuità tra le varie funzioni vengono definite solo per scopi specifici. Il metodo di selezione dei punti usato per il calcolo della funzione interpolante varia tra i vari metodi implementabili.

#### ***Inverso della distanza (IDW)***

Questo metodo si basa sul presupposto che il valore in un punto non campionato può essere approssimato attraverso una media pesata dei valori dei punti che si trovano al di sotto di una determinata distanza o alternativamente da un dato numero di punti più prossimi. I pesi sono inversamente proporzionali ad una potenza della distanza e quindi dato un punto non campionato  $r$ , lo stimatore è

$$F(r) = \sum_{i=1}^m w_i z(r_i) = \frac{\sum_{i=1}^m z(r_i)/|r - r_i|^p}{\sum_{j=1}^m 1/|r - r_j|^p}$$

dove comunemente  $p=1$  e rappresenta il coefficiente di decadimento ovvero con quale modalità esponenziale (lineare, quadratica ecc..) diminuisce l'influenza dei valori dei punti con la distanza.

Nonostante sia un metodo semplice da implementare ha alcuni limiti che ne riducono l'applicabilità. Questo metodo produce forme che si discostano dall'andamento generale dei dati e produce localmente valori estremi.

### ***Natural neighbor***

Tale metodo impiega delle medie pesate dei dati derivati dai punti vicini ricavati dai poligoni di Thiessen (detti anche di Voronoi). Per funzione trivariate i punti noti sono i centroidi di poliedri di Thiessen. Il valore nei punti non campionati è calcolato come media pesata dei punti prossimi con pesi dipendenti dall'area/volume piuttosto che dalle distanze. Il numero di punti usati per il calcolo dei punti incogniti è variabile e dipende dalla configurazione spaziale del set di dati. La superficie risultante, basandosi sui poligoni di prossimità, generalmente è caratterizzata da bruschi cambi di valore.

### ***L'interpolazione basata sulle TIN (Triangulated Irregular Network)***

La generazione delle TIN realizza una tassellazione della superficie a partire dai punti dati. Tale tassellazione è generata tramite triangolazione dei punti utilizzando specifici criteri quali quello di Delunay che utilizza il metodo dei poligoni di Thiessen per individuare i punti da utilizzare come vertici dei triangoli. La TIN risulta quindi in una struttura a faccette triangolari formata da segmenti e nodi dotati di topologia. La stima dei punti incogniti è perciò realizzata da una funzione bivariata determinata, per ogni triangolo utilizzato,



dalla sua vergenza. Le TIN più utilizzate sono caratterizzate da faccette planari, ma è possibile sfruttare la triangolazione per generare superfici non lineari ed ulteriori criteri di continuità per il primo e secondo ordine di derivate assicurando una connessione continua dei triangoli e differenziabilità della superficie risultante. A causa dell'approccio locale il metodo si rileva rapido e, come già accennato, incorpora facilmente le discontinuità e le caratteristiche geometriche delle strutture. Un'appropriata triangolazione per rispettare la geometria della superficie è un'aspetto cruciale del metodo poiché possono insorgere problemi soprattutto per le TIN generate da linee di livello (Bonin & Rousseaux 2005, Weibel & Brändli 1995). Tramite le TIN è possibile realizzare una efficace rappresentazione delle superfici, utili per varie applicazioni come la visualizzazione dinamica e l'analisi della visibilità. Simili alle TIN sono i metodi basati sui rettangoli che adattano una funzione polinomiale non-lineare a rettangoli regolari od irregolari, ad esempio i metodi Bezier, Hermite, B-spline ecc. spesso provvisti della possibilità di regolare il valore della tensione. Tali metodi sono stati sviluppati per gli ambienti CAD e sono poco utilizzati in ambiente GIS.

#### **3.4.4 Metodi geostatistici : *Il Kriging***

Il Kriging (Isaaks & Srivastava 1989) si basa sul concetto di funzione casuale: la superficie o un volume viene considerato come una realizzazione di una funzione casuale avente una determinata covarianza spaziale. Utilizzando i punti dati  $v_i$  e assumendo stazionarietà si può stimare un semivariogramma (più semplicemente chiamato variogramma) definito come funzione di un valore di separazione  $\mathbf{h}$ :

$$\gamma(\mathbf{h}) = \frac{1}{2N(\mathbf{h})} \sum_{(i,j)|\mathbf{h}_{ij}=\mathbf{h}} (v_i - v_j)^2$$

che è legata alla funzione di covarianza  $\mathbf{C}(\mathbf{h})$  da

dove  $\mathbf{C}(\mathbf{0})$  è il valore del semivariogramma per l'infinito.

$$\gamma(h) = C(0) - C(h)$$

La funzione di covarianza è funzione della distanza di separazione  $\mathbf{h}$  tra i punti dati:

$$C(\mathbf{h}) = \frac{1}{N(\mathbf{h})} \sum_{(i,j)|\mathbf{h}_{ij}=\mathbf{h}} v_i \cdot v_j - m_{-\mathbf{h}} \cdot m_{+\mathbf{h}}$$

dove tutti i punti  $v_i$  e  $v_j$  sono separati da un valore di distanza  $\mathbf{h}$  ed  $m_{-\mathbf{h}}, m_{+\mathbf{h}}$

sono le medie dei valori distanti tra loro  $\mathbf{h}$  :  $m_{-\mathbf{h}} = \sum_{i|\mathbf{h}_{ij}=\mathbf{h}} v_i$  ,  $m_{+\mathbf{h}} = \sum_{j|\mathbf{h}_{ij}=\mathbf{h}} v_i$

La funzione  $\gamma(h)$  può essere calcolata per punti separati dalla distanza  $\mathbf{h}$  lungo direzioni specifiche (variogramma direzionale) oppure per tutti i punti del set di dati, separati tra loro da tale distanza, indipendentemente dalla direzione (variogramma omnidirezionale). A seconda della direzione la forma del variogramma può essere differente ossia i dati possono evidenziare una struttura anisotropica. Per dati isotropici, il variogramma può essere semplificato da una funzione radiale  $F(\mathbf{r})$  dipendente dal valore assoluto della distanza di separazione,  $|\mathbf{h}|$ . Esistono varie funzioni che possono essere usate per la modellizzazione di un variogramma (sferiche, esponenziali, gaussiane, Bassel ecc.). I parametri di tali funzioni vengono ottimizzati per adattarsi il più possibile ai valori ottenuti dal variogramma ottenuti dai punti dati.

La superficie interpolata viene quindi costruita utilizzando condizioni statistiche di affidabilità e minima varianza. L'interpolante universal Kriging, è definito come:

$$F(\mathbf{r}) = T(\mathbf{r}) + \sum_{j=1}^N \lambda_j C(\mathbf{r} - \mathbf{r}_j)$$

dove  $T(\mathbf{r})$  rappresenta la componente non casuale espressa come una combinazione lineare di monomi di basso ordine. I monomi e i coefficienti  $\lambda_j$  sono individuati risolvendo un sistema di equazioni lineari. In generale il kriging

predice valori in punti e blocchi in uno spazio d-dimensionale e permette di incorporare l'isotropia. Diverse estensioni permettono di migliorarne la flessibilità ed applicabilità. Il Co-Kriging include le informazioni sulla correlazione di due o più attributi per migliorare la qualità dell'interpolazione mentre il Kriging disgiuntivo è utilizzato in applicazioni dove la probabilità che i valori misurati eccedano certi valori limite di interesse. Per i casi in cui la stazionarietà non è considerata valida come assunto di base, si utilizza il Kriging zonale. La principale caratteristica del Kriging è la qualità della previsione e la capacità di prevedere la distribuzione spaziale dell'incertezza. Nonostante ciò ha avuto poco successo nelle applicazioni dove la geometria locale delle superfici e la regolarità delle stesse costituisce un punto cruciale, dove altri metodi si sono rilevati di maggiore utilità.

### **3.5 Parametri morfometrici estraibili dall'analisi differenziale**

L'analisi della superficie terrestre dal punto di vista di una superficie continua è stata oggetto di studio da parte di vari autori. Evans (1972) che ha coniato il termine di “geomorfometria generale” per indicare la disciplina che si occupa di tale tipo di analisi. Mitsova & Hofierka (1993), Shary et alii (2002) offrono una descrizione analitica dei parametri necessari a descrivere la forma della superficie. Gauss (1827) è stato il primo a studiare in maniera differenziale le superfici per formulare le sue teorie sui flussi dei campi elettrici e magnetici. La superficie terrestre è attraversata anch'essa da diversi campi vettoriali, il cui principale è ovviamente il campo gravitazionale. Altri campi sono quello magnetico, che ovviamente non riveste particolare importanza nell'ambito dello studio geomorfologico, e quello legato all'irraggiamento solare, che invece si rivela utile per caratterizzare la superficie dal punto di vista morfodinamico, poiché l'energia solare influenza l'evoluzione del versante anche in maniera indiretta, ad esempio tramite il controllo legato allo sviluppo di differente tipo di vegetazione. Tuttavia la direzione di tale campo e la sua intensità variano sia durante la giornata sia durante l'anno. Tale parametro non verrà dunque trattato di seguito, considerato che si rivela di maggior interesse l'analisi della forma della superficie. Possiamo quindi considerare l'insieme “superficie terrestre+campo gravitazionale” per l'analisi dei parametri morfometrici. Dalla geometria differenziale della superficie emerge l'esistenza di due principali classi di parametri morfometrici: quelli dipendenti dalla direzione del campo vettoriale che attraversa la superficie, e quelli che non vengono influenzati dalla sua direzione e quindi dalla presenza dello stesso. Shary et alii (2002) introducono anche il concetto di varianza ed invarianza rispetto alla rotazione lungo l'asse verticale e l'inversione della superficie (inversione intesa come ribaltamento, ossia una rotazione di 180 gradi lungo uno degli assi orizzontali) che qui però non verrà presa in considerazione.

I principali parametri morfometrici sono ottenuti dalle derivate prime della superficie, che sono la pendenza e l'aspetto e dalle derivate seconde da cui è possibile estrarre vari tipi di curvatura. La geometria della superficie può essere analizzata in maniera soddisfacente attraverso l'analisi di una funzione bivariata  $z=f(x,y)$  continua, ottenuta tramite qualche metodo di interpolazione, fino appunto alle derivate di secondo ordine ed estraendo i parametri topografici tramite le derivate di questa funzione. Quest'approccio si è dimostrato valido dal punto di vista computazionale (Wood 1996) e consente di estrarre pendenza, esposizione, curvatura del profilo, curvatura planare, curvatura tangenziale e le altre curvature.

Introdotte le seguenti abbreviazioni matematiche (Mitasova & Hofierka 1993):

$$f_x = \frac{\partial z}{\partial x} \quad f_y = \frac{\partial z}{\partial y} \quad f_{xx} = \frac{\partial^2 z}{\partial x^2} \quad f_{yy} = \frac{\partial^2 z}{\partial y^2} \quad f_{xy} = \frac{\partial^2 z}{\partial x \partial y}$$

$$p = \sqrt{f_x^2 + f_y^2} \quad q = p + l$$

l'angolo di massima pendenza  $\gamma$  e l'angolo dell'esposizione  $\alpha$  sono calcolati a partire dal gradiente della funzione  $\text{Grad}(f)=(f_x, f_y)$  come segue:

$$\gamma = \arctan(p) \quad \alpha = \arctan\left(\frac{f_y}{f_x}\right)$$

Il calcolo della curvatura risulta più laborioso poiché in genere una superficie presenta curvature differenti per ogni direzione, e la curvatura da calcolare va scelta in base al tipo di applicazione che se ne intende fare. Nel campo delle Scienze della Terra la curvatura nella direzione del gradiente (curvatura del profilo) risulta di primaria importanza poiché riflette il cambiamento dell'angolo di pendenza e quindi determina una variazione nella velocità dei flussi idrici lungo il pendio. La curvatura in una direzione perpendicolare al gradiente riflette le variazioni dell'angolo dell'aspetto ed influenza la divergenza/convergenza delle linee di flusso idrico. Tale curvatura è normalmente calcolata lungo il piano orizzontale come curvatura delle linee di livello ed è chiamata curvatura planare,

L'analisi della superficie terrestre dal punto di vista di una superficie continua è stata oggetto di studio da parte di vari autori. Evans (1972) che ha coniato il termine di “geomorfometria generale” per indicare la disciplina che si occupa di tale tipo di analisi. Mitsova & Hofierka (1993), Shary et alii (2002) offrono una descrizione analitica dei parametri necessari a descrivere la forma della superficie. Gauss (1827) è stato il primo a studiare in maniera differenziale le superfici per formulare le sue teorie sui flussi dei campi elettrici e magnetici. La superficie terrestre è attraversata anch'essa da diversi campi vettoriali, il cui principale è ovviamente il campo gravitazionale. Altri campi sono quello magnetico, che ovviamente non riveste particolare importanza nell'ambito dello studio geomorfologico, e quello legato all'irraggiamento solare, che invece si rivela utile per caratterizzare la superficie dal punto di vista morfodinamico, poiché l'energia solare influenza l'evoluzione del versante anche in maniera indiretta, ad esempio tramite il controllo legato allo sviluppo di differente tipo di vegetazione. Tuttavia la direzione di tale campo e la sua intensità variano sia durante la giornata sia durante l'anno. Tale parametro non verrà dunque trattato di seguito, considerato che si rivela di maggior interesse l'analisi della forma della superficie. Possiamo quindi considerare l'insieme “superficie terrestre+campo gravitazionale” per l'analisi dei parametri morfometrici. Dalla geometria differenziale della superficie emerge l'esistenza di due principali classi di parametri morfometrici: quelli dipendenti dalla direzione del campo vettoriale che attraversa la superficie, e quelli che non vengono influenzati dalla sua direzione e quindi dalla presenza dello stesso. Shary et alii (2002) introducono anche il concetto di varianza ed invarianza rispetto alla rotazione lungo l'asse verticale e l'inversione della superficie (inversione intesa come ribaltamento, ossia una rotazione di 180 gradi lungo uno degli assi orizzontali) che qui però non verrà presa in considerazione.

I principali parametri morfometrici sono ottenuti dalle derivate prime della superficie, che sono la pendenza e l'aspetto e dalle derivate seconde da cui è

possibile estrarre vari tipi di curvatura. La geometria della superficie può essere analizzata in maniera soddisfacente attraverso l'analisi di una funzione bivariata  $z=f(x,y)$  continua, ottenuta tramite qualche metodo di interpolazione, fino appunto alle derivate di secondo ordine ed estraendo i parametri topografici tramite le derivate di questa funzione. Quest'approccio si è dimostrato valido dal punto di vista computazionale (Wood 1996) e consente di estrarre pendenza, esposizione, curvatura del profilo, curvatura planare, curvatura tangenziale e le altre curvature.

Introdotte le seguenti abbreviazioni matematiche (Mitasova & Hofierka 1993):

$$f_x = \frac{\partial z}{\partial x} \quad f_y = \frac{\partial z}{\partial y} \quad f_{xx} = \frac{\partial^2 z}{\partial x^2} \quad f_{yy} = \frac{\partial^2 z}{\partial y^2} \quad f_{xy} = \frac{\partial^2 z}{\partial x \partial y}$$

$$p = \sqrt{f_x^2 + f_y^2} \quad q = p + l$$

l'angolo di massima pendenza  $\gamma$  e l'angolo dell'esposizione  $\alpha$  sono calcolati a partire dal gradiente della funzione  $\text{Grad}(f) = (f_x, f_y)$  come segue:

$$\gamma = \arctan(p) \quad \alpha = \arctan\left(\frac{f_y}{f_x}\right)$$

Il calcolo della curvatura risulta più laborioso poiché in genere una superficie presenta curvature differenti per ogni direzione, e la curvatura da calcolare va scelta in base al tipo di applicazione che se ne intende fare. Nel campo delle Scienze della Terra la curvatura nella direzione del gradiente (curvatura del profilo) risulta di primaria importanza poiché riflette il cambiamento dell'angolo di pendenza e quindi determina una variazione nella velocità dei flussi idrici lungo il pendio. La curvatura in una direzione perpendicolare al gradiente riflette le variazioni dell'angolo dell'aspetto ed influenza la divergenza/convergenza delle linee di flusso idrico. Tale curvatura è normalmente calcolata lungo il piano orizzontale come curvatura delle linee di livello ed è chiamata curvatura planare,

un altro tipo di curvatura calcolata lungo una direzione tangente alle isoipse ma in un piano normale alla superficie viene denominata curvatura tangenziale. Le equazioni per il calcolo delle curvature possono essere derivate dall'equazione generale della curvatura  $\mathbf{k}$  della sezione di un piano che attraversa la superficie in un dato punto:

$$\mathbf{k} = \frac{f_{xx} \cos^2 \beta_1 + 2 f_{xy} \cos \beta_1 \cos \beta_2 + f_{yy} \cos^2 \beta_2}{\sqrt{q} \cos \theta}$$

dove  $\theta$  è l'angolo tra la normale della superficie in un punto dato ed il piano che taglia la superficie;  $\beta_1, \beta_2$  sono gli angoli tra la tangente della sezione e gli assi x,y rispettivamente.

L'equazione per la curvatura del profilo  $\mathbf{k}_s$  è calcolata come la curvatura della sezione di un piano normale nella direzione del gradiente:

$$\mathbf{k}_s = \frac{f_{xx} f_x^2 + 2 f_{xy} f_x f_y + f_{yy} f_y^2}{p \sqrt{q^3}}$$

con

$$\cos \beta_1 = \frac{f_x}{\sqrt{(pq)}} \quad \cos \beta_2 = \frac{f_y}{\sqrt{(pq)}} \quad \cos \theta = 1$$

similmente l'equazione per la curvatura planare  $\mathbf{k}_h$  è la curvatura di una sezione di un piano orizzontale (parallelo al piano x-y):

$$\mathbf{k}_h = \frac{f_{xx} f_y^2 + 2 f_{xy} f_x f_y + f_{yy} f_x^2}{\sqrt{p^3}}$$

ed

$$\cos \beta_1 = \frac{f_y}{\sqrt{(p)}} \quad \cos \beta_2 = -\frac{f_x}{\sqrt{(p)}} \quad \cos \theta = \sqrt{\frac{p}{q}}$$

L'equazione per la curvatura tangenziale  $\mathbf{k}_t$ , è calcolata come la curvatura di una sezione di un piano normale nella direzione perpendicolare al gradiente:

$$\mathbf{k}_t = \frac{f_{xx} f_y^2 - 2 f_{xy} f_x f_y + f_{yy} f_x^2}{p \sqrt{q}}$$



con

$$\cos \beta_1 = \frac{f_y}{\sqrt{(p)}} \quad \cos \beta_2 = -\frac{f_x}{\sqrt{(p)}} \quad \cos \theta = 1$$

Due curvature importanti sono la curvatura massima, **kmax**, e la curvatura minima, **kmin**, che possono essere calcolate a partire da altre curvature. La media di due qualsiasi curvature estratte nel punto X da due sezioni ottenute dall'intersezione di due piani ortogonali non cambia ruotando la direzione dell'insieme dei piani. Tale curvatura H, chiamata appunto curvatura media, può essere estratta da qualsiasi coppia di curvature ortogonali.

Ad esempio  $H=1/2 (k_t+k_s)=1/2(k_{min}+k_{max})$ . La formula generale è la seguente:

$$H = -[(1 + f_y^2) f_{xx} - 2 f_x f_y f_{xy} + (1 + f_x^2) f_{yy}] / [2(1 + f_x^2 + f_y^2)^{3/2}]$$

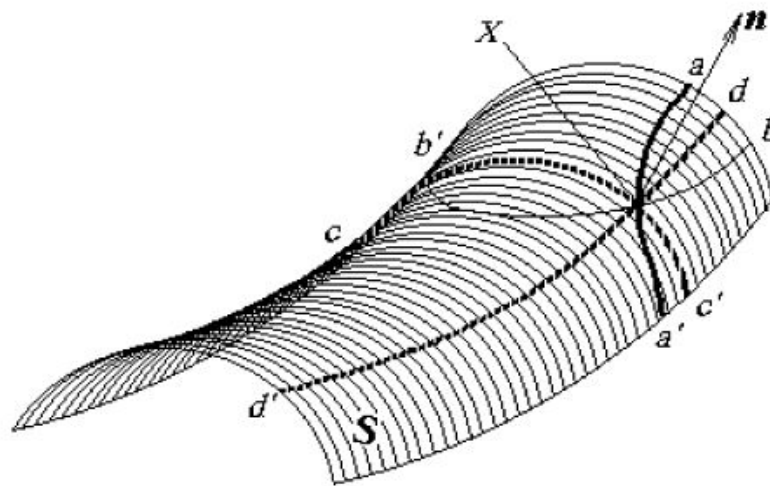
Le curvatura massima e minima si possono ottenere dalla curvatura media e dal coefficiente di asfericità M. Tale parametro identifica di quanto si discosta una superficie dalla forma sferica ed equivale a zero per una sfera perfetta:

$$M = [\{ f_{xx} [(1 + f_y^2)/(1 + f_x^2)]^{1/2} - f_{yy} / [(1 + f_y^2)/(1 + f_x^2)]^{1/2} \}^2 / (1 + f_x^2 + f_y^2) + \{ f_x f_y f_{xx} [(1 + f_y^2)/(1 + f_x^2)]^{1/2} - 2 [(1 + f_y^2)(1 + f_x^2)]^{1/2} f_{xy} + f_x f_y f_{yy} / [(1 + f_y^2)/(1 + f_x^2)]^{1/2} \}^2 / 2]^{1/2} / [2(1 + f_x^2 + f_y^2)^{3/2}]$$

Una volta definiti H ed M le curvature massime e minime possono essere ricavate essendo  $K_{min}=H-M$  ed  $K_{max}=H+M$ . Queste curvature sono utili a descrivere la forma di superfici la cui natura non è determinata principalmente da processi di erosione fluviale o movimenti di massa. Infatti sono curvature indipendenti dal campo gravitazionale. Superfici di tale tipo possono essere ad esempio i versanti di faglia. E' possibile derivare altri tipi di curvature quali il rotore, che identifica la curvatura orizzontale delle linee di flusso. Tutte le altre curvature sono comunque ricavabili dalle formule di quelle principali (Shary et alii 2002).

I valori positivi e negativi delle curvature, i valori di pendenza ed esposizione,

possono essere combinati per definire le forme geometriche di base del rilievo. Ogni forma base comporta una combinazione differente dei valori dei parametri morfometrici qui introdotti. Per utilizzare tale informazione in un sistema di classificazione automatica, questi parametri vengono associati a vettori che costituiscono gli attributi della superficie topografica in uno spazio multivariato.



*Fig. 3.5 La sezione che scaturisce dall'intersezione di un piano e superficie la superficie  $S$  determina la curvatura nel punto  $X$ . A seconda della direzione di immersione del piano emergono differenti tipi di curvatura.  $n$  è il vettore normale alla superficie al punto  $X$ .  $cc'$  è la sezione che determina la curvatura massima,  $dd'$  la sezione che determina la curvatura minima,  $bb'$  è la sezione della curvatura del profilo e  $aa'$  la sezione lungo la quale emerge la curvatura tangenziale. (da Shary et alii 2002)*

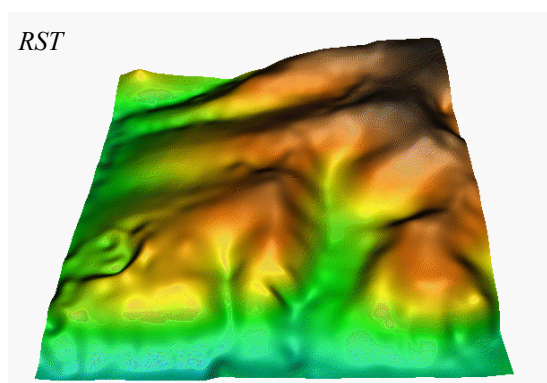
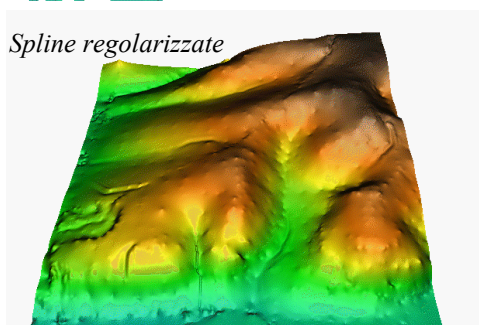
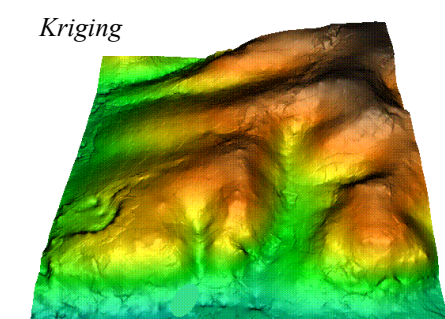
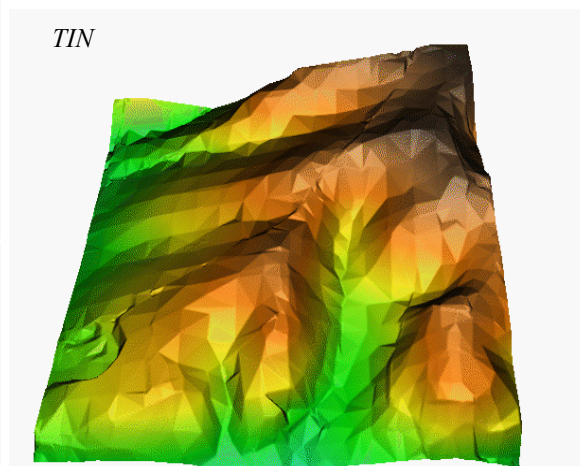
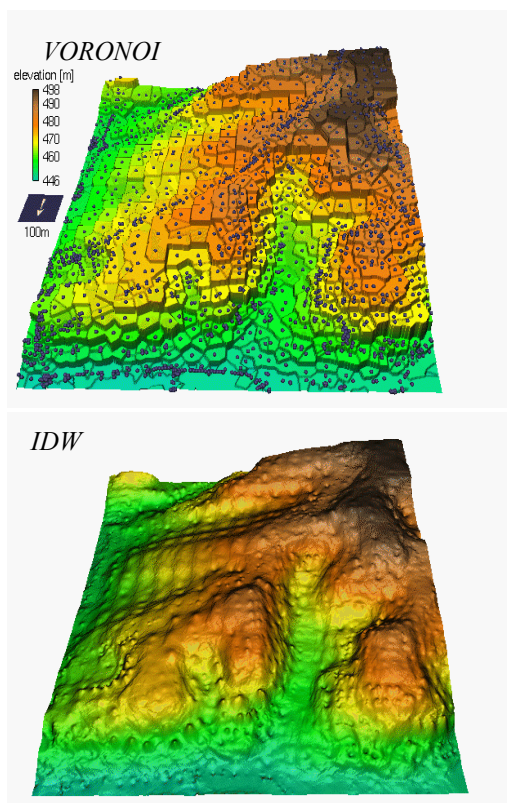


Fig 3.6 Superficie risultante da vari metodi di interpolazione ottenuti a partire dallo stesso set di dati (da Mitasova & Mitas 1999)

### 3.6 Parametri determinati dal flusso idrico

Il campo gravitazionale che agisce in congiunzione con la forma del rilievo produce un gradiente che agisce sui flussi idrici. L'acqua di superficie segue un percorso determinato dalla direzione del gradiente e i flussi convergenti e divergenti lungo la superficie sono determinati dai parametri descritti precedentemente. Tali parametri risultano però insufficienti a descrivere l'andamento globale dei flussi poiché si limitano a descrivere la geometria locale della superficie. In questo paragrafo si descrivono le metodologie note in letteratura che sono alla base della simulazione del flusso superficiale tramite DTM e dei principali parametri topografici derivati. Essendo il flusso idrico determinato dalla forma del rilievo, i parametri derivati si rivelano molto utili per la classificazione automatica. Mentre i parametri morfometrici visti in precedenza sono valori estratti da un punto della superficie interpolata, i parametri legati al flusso idrico derivano da valori cumulati, la cui proprietà dipende da vari fattori quali l'algoritmo utilizzato per il calcolo del flusso e la forma della superficie.

Il calcolo del flusso, determinato mediante gli algoritmi integrati comunemente nei GIS, si basa sull'assunto che la superficie terrestre rappresentata dal DTM subisce un *runoff* omogeneo, ovvero che esista una condizione di *steady state*. In pratica, per simulare tale condizione, ad ogni cella è associato un valore di contributo al flusso pari ad 1. Si presuppone inoltre che non vi sia né infiltrazione nel sottosuolo, né evapotraspirazione, sicché i parametri sono determinati solo dalla forma della superficie e da come l'algoritmo distribuisce tale contributo unitario lungo le celle adiacenti. Si tratta di algoritmi semplici rispetto ai modelli di simulazione noti in letteratura (es. Bjoreberg et alii 1997, Moore et alii 1991), . Comunque alcuni parametri estraibili da tali algoritmi sono anche utilizzabili all'interno di modelli di erosione e sedimentazione quali il RUSLE (Mitasosva et alii 1996a). Ad esempio si possono estrarre i valori di *slope lenght* (LS) e *slope steepness* (S) utilizzati da tali modelli. Lo scopo principale di questi algoritmi rimane comunque quello della

delineazione automatica del reticolo idrografico e la suddivisione dell'area in bacini idrografici. Da questo punto di vista tutti gli algoritmi utilizzano, con alcune varianti, metodi che calcolano come il contributo idrico si distribuisce a partire da una cella del *raster*, lungo le celle adiacenti. Essendo la forza di gravità il motore del flusso idrico, le celle vengono processate partendo da quella posta a quota più elevata e procedendo il calcolo verso il basso. Nel *raster* che rappresenta una porzione di superficie terrestre ben variegata si trovano molte celle che rappresentano la cima di un rilievo. Questo fatto consente di poter processare porzioni di DTM in maniera indipendente poiché è escluso da tali algoritmi il flusso idrico contro la forza di gravità. Risulta quindi possibile parallelizzare il calcolo rendendolo più rapido. Ciò si rivela particolarmente utile nel caso si voglia elaborare un DTM di grandi dimensioni. Le differenti tecniche differiscono principalmente nel modello di deflusso rispetto alla geometria del rilievo. Nello specifico, le principali differenze si riscontrano in come il flusso è distribuito lungo le celle adiacenti. I più semplici modelli considerando il percorso del flusso idrico di tipo lineare, cioè da un punto ad un altro della matrice. Ciò significa che una cella del *raster* contribuisce ad alimentare una sola cella del *raster* tra quelle adiacenti.

Nel determinare il flusso si possono individuare problemi dovuti al fatto che il modello digitale del terreno non è stato generato in maniera tale da includere i vincoli altimetrici determinati dalla presenza del reale reticolo idrografico o comunque problemi dovuti alla presenza di difetti prodotti dal processo di interpolazione. In effetti uno dei principali vantaggi nell'utilizzare un DTM per l'analisi del reticolo idrografico è proprio quello di poter estrarre lo stesso in maniera automatica. Questo è un vantaggio notevole in termini di tempo poiché includere il reticolo idrografico presuppone il doverlo digitalizzare come elemento vettoriale. Un'altra considerazione è che difatti l'informazione sul reticolo idrografico è disponibile solamente se la fonte lo include. Ad esempio, se il DTM è derivato invece che da tavolette topografiche, da una scansione LIDAR

o da interferometria radar si può ottenere solamente il dato altimetrico campionato o già interpolato. Se si utilizza una tavoletta topografica il reticolo idrografico è rappresentato dalle “linee blu”. Queste linee rappresentano il reale reticolo idrografico come è stato rilevato dai topografi. Se si ha il tempo di digitalizzare tale reticolo è possibile evitare l'estrazione automatica da DTM. Tale reticolo risulta più preciso rispetto a quello ottenibile in automatico. Tuttavia non essendo dotato di quota è utilizzabile solo per analizzare la sua geometria dal punto di vista planimetrico. Sovrapponendo tale reticolo ad un DTM interpolato dalle linee di livello è invece possibile analizzarlo anche nelle sue caratteristiche altimetriche. Inoltre è possibile utilizzare il reticolo vettorializzato come vincolo per rendere il DTM "idrologicamente corretto". Ad esempio esistono software che generano TIN e che possono utilizzare il reticolo per guidare il processo di triangolazione tramite l'inclusione delle cosiddette *breaklines* (ArcView 1993). Altri software quali ANUDEM (Hutchinson 1989) possono interpolare direttamente il raster da linee di livello imponendo dei vincoli lungo il reticolo idrografico e le crenulazioni. Dove non è possibile creare ad hoc modelli digitali idrologicamente corretti, perché non si ha a disposizione software adatto o si utilizzano modelli forniti del solo dato altimetrico, è necessario correggere a posteriori tramite specifiche routine. Le aree problematiche nei modelli *raster* sono ad esempio i *sink* isolati, ovvero depressioni formate da una singola cella. Per depressione si intende una cella caratterizzata da una quota inferiore rispetto a tutte le 8 celle adiacenti. Oltre ai *sink* isolati esistono le aree piatte. Queste sono aree formate da più celle adiacenti aventi quota identica. E' evidente che in tali aree non è possibile calcolare un valore di direzione e quindi il calcolo della direzione di flusso risulta particolarmente problematico. Bisogna distinguere tra aree piatte aventi un punto di deflusso ed aree che non posseggono nessun punto di deflusso. Le prime sono caratterizzate da avere, tra le celle che bordano l'area, almeno una cella, evidentemente a quota più bassa, che permette il deflusso del reticolo verso le altre celle del *raster*. Un'area piatta estesa che non possiede un

punto di deflusso è un *sink* in grado di catturare una grande quantità di flusso idrico, come un lago. Le aree piatte sono particolarmente problematiche a causa della mancanza di informazione sulla direzione del flusso per diverse celle adiacenti, per cui sono necessari dei metodi particolari per definire il flusso lungo queste aree problematiche. Sia i metodi basati sulla direzione singola che quelli basati sulla direzione multipla del flusso, non funzionano se non è definita per ogni cella un valore di direzione. Un metodo molto diffuso per assegnare la direzione del flusso, introdotto da Jenson & Domingue (1988), si basa sul riempimento di tali aree in maniera iterativa, fino a trovare una via di fuga del flusso idrico. Per le aree piatte che posseggono già una cella per cui è possibile calcolare la direzione di deflusso tale riempimento va semplicemente ad individuare tale cella, se ve ne è più di una il deflusso viene calcolato lungo tutte queste celle. Nel caso delle depressioni isolate e delle depressioni piatte estese, l'area viene riempita, aumentando la quota delle celle, fino a che tale area in effetti non si raggiunge la quota di una cella per cui è possibile, ricordando sempre che il deflusso è definito solo se una cella ha almeno una cella adiacente posta a quota più bassa, calcolare una direzione di deflusso. La direzione del flusso per ogni cella delle aree piatte viene quindi calcolata iterativamente partendo dai bordi esterni delle aree piatte (evidentemente più elevati) al punto di deflusso. Altri metodi sono stati proposti da Martz et Garbrecht (1999) e Jones (2002). I metodi più diffusi basati sul calcolo della direzione di flusso sono il D8 da deterministic 8 (O'Callaghan & Marks 1984) che distribuisce tutto il contributo idrico sulla cella altimetricamente più bassa tra quelle adiacenti, il Metodo del *Multiple Flow Direction* o MFD (Quinn et alii 1991) che distribuisce il contributo su più celle in maniera proporzionale al valore di aspetto, ed il Metodo D-infinito (Tarboton 1997) simile al D8, in quanto il contributo idrico è distribuito ad un sola cella, ma calcola non in base alla differenza altimetrica delle celle adiacenti, ma in base all'aspetto, che assume valori continui tra 0 e 360 gradi (da qui il termine direzione infinita). Tutti questi

metodi determinano la possibilità di calcolare vari parametri il cui valore dipende da come si distribuisce il flusso idrico e dall'algoritmo utilizzato. Il più immediato è l'area di cattura (CA), che rappresenta l'area che contribuisce al deflusso idrico di una determinata cella, o altri composti come il *wetness index* o lo *stream power* che rappresentano rispettivamente la saturazione e il potere erosivo del flusso idrico in un determinato punto della superficie, sempre in condizioni di *steady state* (Wilson & Gallant 2000).

### 3.7 Qualità dei DTM

La qualità di un DTM dipende principalmente dalla densità e accuratezza di campionamento e dal metodo di interpolazione. Nel caso del metodo di interpolazione molti esempi di tecniche utilizzate in letteratura, anche abbastanza recente, descrivono modelli lineari, o l'interpolazione bilineare (Kidner et alii 1999) o le già citate TIN per la generazione di DTM. Questi modelli possono produrre una serie diversa di soluzioni a partire dal medesimo campionamento. I modelli di tipo lineare nonostante la loro limitata resa qualitativa sono molto usati per la loro semplicità computazionale. Tenendo presente che per risolvere i modelli lineari sono necessari soli tre punti vicini (o quattro per l'interpolazione bilineare) appare ovvio come da un interpolante di ordine superiore, che richiede un maggior numero di punti vicini per poter risolvere il relativo sistema di equazioni, si otterranno risultati superiori rispetto a qualsiasi modello lineare. Un tempo la limitata potenza di calcolo era un limite insuperabile e quindi i metodi di interpolazione più sofisticati erano in grado di risolvere in tempi ragionevoli sistemi costituiti al massimo da poche migliaia di punti mentre un modello digitale del terreno per poter essere generato necessita di decine o centinaia di migliaia di punti campione. Tuttavia la potenza dei normali computer domestici è cresciuta ed oggi computer in grado di eseguire miliardi di operazioni al secondo



sono la norma; per cui anche sistemi costituiti da centinaia di migliaia di punti possono essere risolti in tempi ragionevoli. Ciò ha portato all'inclusione di tali metodi nei pacchetti GIS ed il loro utilizzo per la generazione di DTM.

Un altro problema può essere causato dall'inclusione di un numero eccessivo di punti. Il terreno può esibire sia andamenti molto generali che fluttuazioni locali con componenti in frequenza rispettivamente bassi ed alti. Risulta quindi molto importante considerare il numero di punti ed i parametri di interpolazione da cui è possibile ottenere i migliori risultati. Per quanto riguarda i metodi di stima della qualità dei DTM, un metodo di accuratezza globale consiste nel misurare l'errore quadratico medio (Wise 1998)

$$\sqrt{\frac{\sum (Z_{est} - Z_{obs})^2}{n}}$$

dove  $Z_{est}$  e  $Z_{obs}$  sono le differenze di quota tra due punti del DTM interpolato e la superficie di riferimento o un sottoinsieme di punti campionati, non utilizzati durante il processo di interpolazione. Più importante dell'accuratezza globale del dato interpolato, risulta l'accuratezza locale nel riprodurre la forma della superficie. Ad esempio, confrontando il risultato di due interpolazioni dello stesso DTM, quello caratterizzato da un'errore quadratico medio più basso può presentare difetti più evidenti rispetto al secondo. Un tipo di difetto spesso presente nei DTM interpolati, e che affligge pesantemente l'accuratezza nel calcolo delle derivate della superficie, è il cosiddetto “waving” o “terracing” ed il

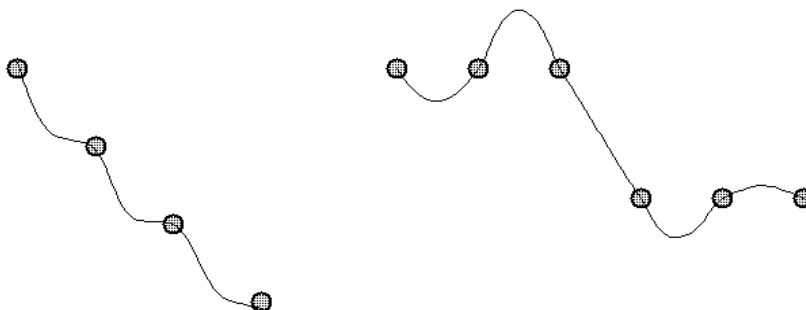
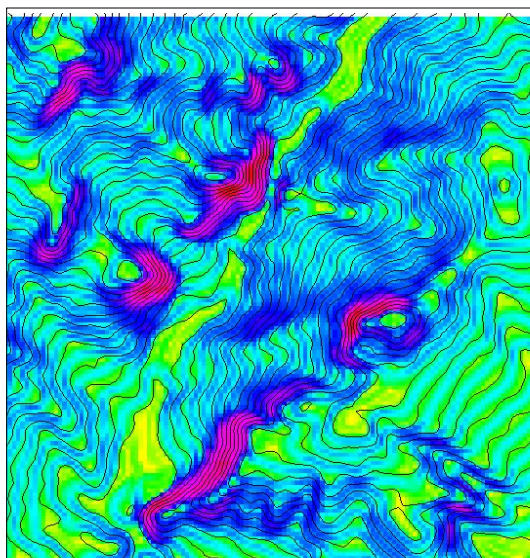
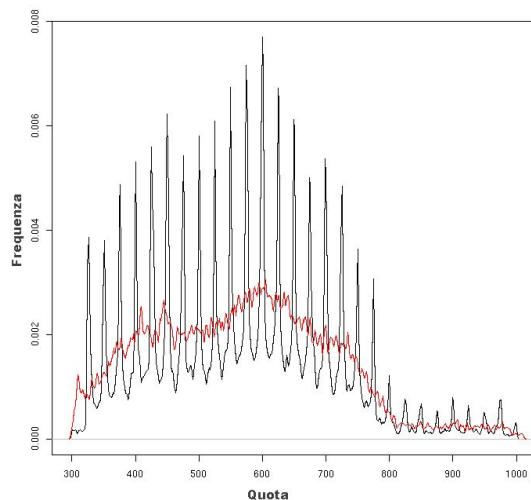


Fig 3.7 Tipici errori di interpolazione: A sinistra "terracing", a destra "ringing" (da Frankie & Gousie <http://www.ecse.rpi.edu/Homepages/wrf/>)

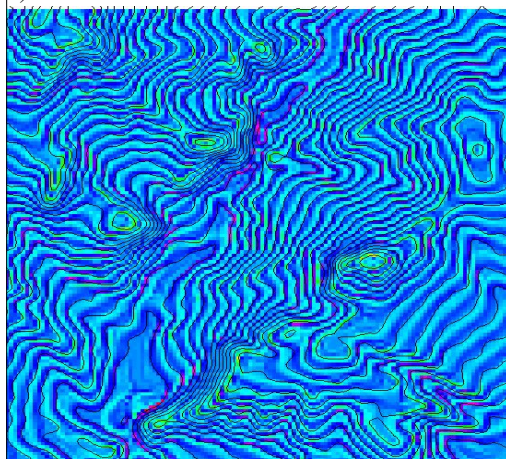
“ringing” (Frankie & Gousie). Un certo grado di “waving” e “ringing” è presente nel risultato di praticamente tutti gli algoritmi di interpolazione noti in letteratura. Ad esempio, nel caso delle spline, il “waving” è dovuto al fatto che la funzione è vincolata ai punti tra le isoipse e l'energia di deformazione della superficie vincolata produce un andamento concavo-convesso tra un'isoipsa e l'altra. Il “ringing” si produce dove due isoipse consecutive hanno la stessa quota ed è dovuta sempre al comportamento di tipo elastico della funzione. Questi difetti si amplificano nel calcolo delle derivate della superficie. La presenza del “waving” è facilmente riscontrabile sia attraverso l'analisi visiva delle mappe a falsi colori delle derivate della superficie, sia attraverso l'analisi dell'istogramma della frequenza di distribuzione della quota. L'istogramma (Fig 3.8c) presenta picchi ad intervalli regolari di ampiezza pari alla equidistanza delle isoipse utilizzate per generare il DTM. Questi difetti possono essere attenuati in vari modi. Nel caso il DTM sia già stato interpolato è possibile utilizzare una tecnica di *smoothing* del *raster* utilizzando un filtro passa basso che elimini dal DTM le frequenze alle quali si presentano tali difetti. Se l'algoritmo interpolante lo permette, è possibile ridurre il problema tramite una opportuna scelta dei parametri di interpolazione (cfr. 3.4.1). In certi casi è possibile modificare i parametri dell'algoritmo di interpolazione; ad esempio diminuendo il valore di tensione ed aumentando il valore di smoothing (fig. 3.8d,e) del metodo di interpolazione RST, si riduce notevolmente l'effetto di terracing e di ringing (Cebecauer 2002, Mitsova & Hofierka 1993).



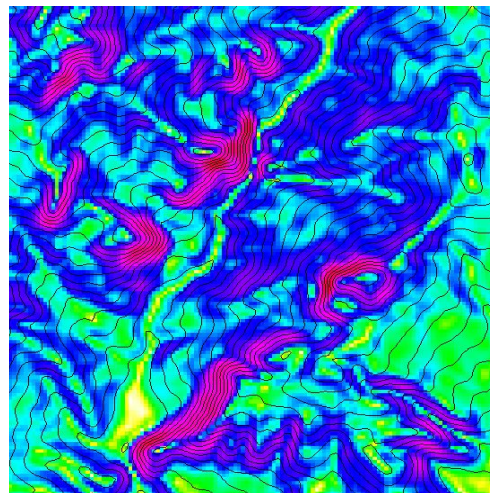
A)



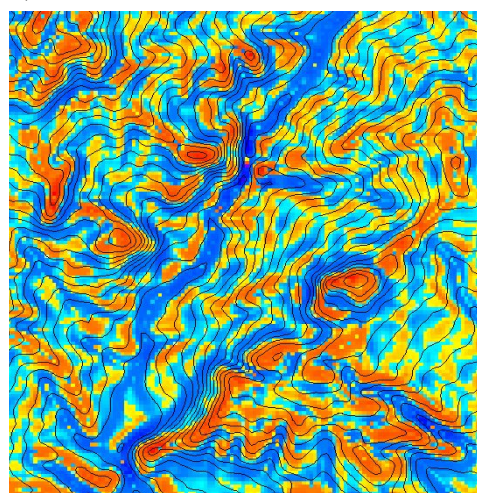
C)



B)



D)



E)

*Fig. 3.8 – Analisi del DTM della valle del F. Volturno*

*A) Raster della pendenza prima dello smoothing che evidenzia l'effetto di terracing*

*B) Raster della curvatura del profilo (derivata seconda) prima dello smoothing: il terracing risulta molto più evidente*

*C) Funzione di densità della quota del DTM. In nero prima dello smoothing, in rosso dopo lo smoothing, i picchi risultano notevolmente ridotti. I picchi della curva in nero presentano la medesima equidistanziana delle isoipse da cui è stato interpolato il DTM*

*D) Raster della pendenza dopo lo smoothing*

*E) Raster della curvatura del profilo dopo lo smoothing*

## B.III Operazioni sui raster

### 4.1 Il modello raster

Come descritto nel capitolo 3, tra i modelli di maggior utilizzo nel campo della modellazione della superficie terrestre i modelli di tipo *raster* sono tra i più utilizzati. Le motivazioni sono di ordine pratico. Poiché i modelli del terreno di tipo *raster* posseggono una struttura dati analoga alla struttura delle immagini digitali, è possibile adottare le tecniche di elaborazione sviluppate per il processamento di immagini digitali anche per l'analisi dell'informazione contenuta nel modello digitale del terreno. Rispetto ai modelli vettoriali, quali i *Triangular Irregular Network*, è possibile effettuare analisi spaziale sfruttando la struttura discreta ed ordinata della matrice dati. Un esempio immediato è l'uso dei filtri di convoluzione di cui si discute in un prossimo paragrafo. Il modello *raster* si presta bene all'analisi numerica. In particolare tutta una serie di tecniche di analisi della superficie ed analisi del pattern possono essere implementate tramite algoritmi dalla struttura lineare. Nel caso di altri modelli quali le TIN, è invece necessario utilizzare complessi algoritmi vettoriali per l'analisi geometrica. Risulta difatti complesso differenziare la superficie a causa della struttura a nodi irregolarmente spaziat, e conseguentemente estrarre parametri quali le curvature. Per l'analisi del pattern è più semplice convertire la TIN in un modello *raster* o in un modello a punti regolarmente spaziat.

Nell'ambito dell'analisi *raster* applicata ai modelli digitali del terreno, le più comuni tecniche di elaborazione sono la *Neighborhood analysis* o analisi di vicinanza, la analisi zonale e la convoluzione. Benché gli algoritmi che implementano le varie metodologie siano sostanzialmente procedure che si occupano di processare dei pixel (o celle), sia l'approccio che il contesto nel quale sono nate le metodologie sono sostanzialmente differenti. Ad esempio sia la "Neighborhood analysis" che la convoluzione utilizzano entrambi il concetto di cella focale, ma il tipo di analisi e che si implementa è differente. La neighborhood

analysis serve principalmente per estrarre dei parametri statistici, così pure la analisi zonale, e va quindi inserita nel contesto della statistica spaziale. I filtri di convoluzione nascono nel contesto della teoria dei segnali e trattano il *raster* come un segnale discreto a due dimensioni. Un'immagine *raster* può difatti essere trattata come un segnale bidimensionale del quale si possono analizzare le frequenze di cui è composto tramite trasformata di Fourier (Walker 1988).

Un limite comune a molti pacchetti GIS che gestiscono DTM in formato *raster* è costituito dal fatto che gli algoritmi implementati, soprattutto quelli utilizzati per l'analisi del pattern, gestiscono i livelli informativi con un precisione di 8 bit. Questo limita la gamma di intervalli numerici. Tale limite deriva dal fatto che questi algoritmi sono stati pensati e sviluppati appositamente per l'analisi di immagini satellitari.

Nelle immagini digitali in genere, i valori di intensità luminosa vengono usualmente confinati in un intervallo discreto, generalmente 255 valori, gestibile appunto con appena 8 bit di precisione. Anche le immagini multispettrali fornite dai vari sistemi di rilevamento quali i satelliti della serie Landsat e Spot, che sono composte da più bande, vengono fornite all'utente finale in un formato numerico limitato nell'intervallo dei valori interi. Questo, se risulta adeguato nell'ambito dell'analisi di immagini da satellite, si rivela estremamente limitante nei casi dei modelli digitali del terreno. Basti pensare solamente all'informazione primaria di cui è costituito un DTM. La quota, nelle maggior parte dei casi, varia in un range che va ben oltre il limite imposto dagli 8 bit. Inoltre, in aree dove la differenza altimetrica tra celle adiacenti è molto bassa, come ad esempio in aree di fondovalle, l'informazione altimetrica non può essere rappresentata in maniera adeguata. Si è visto inoltre nel capitolo 3 come sia di fondamentale importanza il poter calcolare le derivate della superficie; un software che in grado di gestire solo valori interi e limitati a pochi bit di precisione non è in grado di elaborare nell'intervallo dei valori continui e quindi utilizzabile per l'analisi differenziale della superficie.

Il problema è dovuto a come la corrispondente funzione continua è rappresentata numericamente all'interno dell'algoritmo di calcolo. Si tratta nello specifico di un *array* bidimensionale  $I(m,n)$  con  $m=0 \dots N-1$  ed  $n=0 \dots M-1$ . Il dato viene rappresentato a video come una matrice o griglia di M righe ed N colonne. Nel caso di immagini, il dato *raster* corrisponde ad un'intensità luminosa quantizzata in un numero  $N_g$  di livelli di grigio, generalmente consistente in una potenza di 2, che definisce la "profondità" dell'informazione. Il dato può ovviamente essere quantizzato utilizzando una profondità qualsiasi. Ad esempio 16 bit corrispondono a 65536 valori interi. Il sistema di calcolo del computer consente di archiviare l'informazione sia come numero intero che come numero in virgola mobile. Per rappresentare numeri reali si utilizza mantissa ed esponente. In tal caso non ha più senso parlare di profondità poiché il dato non è più quantizzato in interi ma corrisponde in valori reali. Si utilizza invece il concetto di precisione che è determinata dal numero di bit attribuiti alla mantissa. Con una precisione di 64 bit è possibile effettuare qualsiasi tipo di analisi della superficie o del pattern senza perdita di informazione poiché ben 11 bit sono associati alla mantissa ed i restanti al segno ed all'esponente che definisce l'intervallo di valori disponibili. In questo modo è possibile effettuare sia analisi differenziale della superficie che applicare ai DTM le tecniche di analisi del pattern descritte nel capitolo 5.

Nell'elaborare un dato *raster* lo scopo è quello di estrarre vari livelli informativi. Il presupposto è che tali livelli possano essere utilizzati per suddividere l'area coperta dal dato *raster* in classi e poter attribuire le celle del *raster* ad una di queste classi utilizzando un sistema di classificazione automatico di tipo supervisionato o di tipo non supervisionato (cfr. capitolo 5).

Lo svantaggio principale di utilizzare un dato *raster* consiste nel gran numero di elementi da elaborare. Nel caso di una semplice matrice di 100x100 celle il dato consiste in 10000 valori, ossia la complessità di elaborazione è proporzionale ad  $N \times M$ .

La situazione diventa ancora più pesante se si comprende come in realtà essendo lo scopo finale la classificazione basata su più caratteristiche estratte dal *raster*, il dato finale sarà di tipo multivariato e quindi gli elementi da elaborare saranno  $N \times M \times K$  con  $K$  numero di livelli informativi estratti.

L'estrazione di nuovi livelli informativi dal *raster* di partenza è anche un'operazione spesso necessaria. Infatti, le variabili che si vogliono utilizzare sono spesso dei descrittori sintetici delle proprietà del *raster* o di un'intorno di pixel, realizzati combinando singole variabili o realizzando una trasformazione del set di dati. Lo scopo è infatti ridurre la quantità di dati da elaborare cercando di ottenere dei descrittori che contengano la più elevata quantità di informazione possibile o comunque che contengano un'informazione tale da rendere possibile, tramite un appropriato algoritmo, di associare ogni singola cella del *raster* ad una classe. Si misurano così nuove variabili a partire dell'informazione contenuta nella struttura del *raster* e che andranno a formare un vettore di attributi o *feature vector* (cfr. paragrafo 5). Un esempio di tali attributi sono i parametri morfometrici estratti dal DTM.

Tale operazione non deve essere confusa con le operazioni volte alla generazione di nuovi attributi, definita *feature extraction* (Webb 2002), che avviene attraverso una combinazione lineare o non lineare dell'insieme dei vettori degli attributi e che consente di ridurre il numero di livelli informativi. Esistono numerosi metodi per generare nuovi attributi a partire da vettori di variabili. Tra i più popolari metodi lineari c'è l'analisi delle componenti principali (Webb, 2002) che consente di ridurre enormemente la mole di dati da elaborare, i metodi basati sull'uso di filtri *raster*, come i filtri wavelet (Walker, 1997) ed i filtri di Gabor (Bovic, 1991), spesso utilizzati per generare un set di dati utile all'analisi multirisoluzione. In questo capitolo discuterà del principio dei filtri di convoluzione, della *neighborhood analysis*, dell'analisi zonale e di altre tecniche di analisi del dato *raster* di fondamentale importanza per l'analisi dei modelli digitali del terreno.

## 4.2 I filtri di convoluzione

I filtri di convoluzione (Pitas 1994, Forsyth & Ponce 2003) sono una classe di filtri molto utilizzati in campo dell'analisi delle immagini e costituiscono la base per l'estrazione di informazioni utili alla classificazione automatica di fotografie e sequenze video digitali. Spesso vengono definiti filtri “box” a causa della forma quadrata della matrice di convoluzione.

Difatti i filtri di convoluzione sono anch'essi dei *raster*. Poiché le operazioni di convoluzione sono operazioni che compiono una serie di trasformazioni tramite matrici, sono facilmente impiegabili anche nell'ambito dell'analisi del dato altimetrico o altro dato derivato ed archiviato in formato *raster*. Infatti, con la diffusione dei DTM in formato *raster*, l'uso dei filtri si è rivelato di grande utilità per estrarre informazioni topografiche (Haralick 1983, Horne 1981). Ad esempio le derivate della superficie possono essere adeguatamente approssimate tramite filtri laplaciani. Il filtro laplaciano può essere applicato per esempio direttamente ad un DTM *raster* per evidenziare interfluvii ed impluvi (Mayer 2000), una sorta di mappa dei crinali e delle depressioni. Grazie alla loro natura *raster*, i filtri di convoluzione possono essere impiegati anche per elaborare dati non altimetrici derivati dai DTM, come ad esempio nel caso degli *shaded-relief* al fine di evidenziare i lineamenti.

Il metodo della convoluzione consiste nel moltiplicare due o più matrici in modo da ottenere un nuovo *raster* tramite una trasformazione di tipo lineare. Le matrici vengono moltiplicate tra loro e quindi il risultato è una combinazione lineare dei loro componenti. In pratica i *raster* vengono considerati come la versione discreta di funzioni scomponibili in frequenze tramite trasformata di Fourier. Un esempio di come opera la convoluzione si può vedere in fig. 4.1.



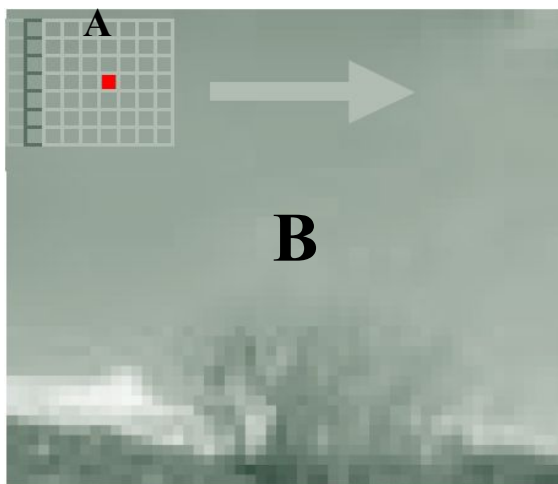


Fig 4.1

Il filtro  $A$  di dimensioni  $8 \times 8$  viene fatto scorrere lungo la matrice  $B$  più grande. Il valore della cella della matrice  $B$  corrispondente alla posizione della cella centrale della matrice  $A$  (in rosso), è detta cella focale, ed assume il valore della somma del prodotto delle due matrici. Poiché tale classe di filtri vengono spesso utilizzati per l'analisi delle immagini, tale valore viene diviso per un fattore tale da farlo ricadere all'interno del range di valori ammessi per la sua codifica (per esempio tra 0 e 255).

Per descrivere la convoluzione tra due segnali si può utilizzare la seguente notazione  $f * g = h$  dove  $g$  è il filtro (o maschera) di convoluzione applicato alla funzione  $f$ , ed  $h$  è il *raster* (o più in generale il segnale) risultante. Per definire la convoluzione tra due *raster* utilizzeremo di seguito la notazione del tipo  $i(m,n)$  che denota l'elemento situato alla riga  $m$  e colonna  $n$  della matrice  $i$ . In tal caso la convoluzione tra due matrici può essere espressa con la seguente notazione:

$$h(i, j) = \sum_{a=Astart}^{a=Aend} \sum_{b=Bstart}^{b=Bend} g(a, b) f(i-a, j-b) \quad (4.1)$$

La convoluzione è un'operazione spesso definita come “point-spread function” ovvero funzione di diffusione del punto. Tale definizione si riferisce al fatto che il valore di un punto, pixel o cella nel caso del *raster*, viene “diffusa” su un certo numero di celle adiacenti, dipendentemente dal tipo e dimensione del filtro

utilizzato. Infatti utilizzando la notazione seguente

$$f(i, j) = \delta(i, j) = \begin{cases} 1 & \text{se } i = j = 0 \\ 0 & \text{altrimenti} \end{cases}$$

essa è rappresentata da un punto circondato da zeri e la convoluzione (4.1) tra le due matrici si riduce a  $h(i, j) = g(i, j)$  cioè il valore del punto centrale è diffuso in tutte le celle della maschera di convoluzione.

La (4.1) corrisponde a far scorrere una finestra mobile, moltiplicare i valori elemento per elemento all'interno della matrice di convoluzione ed associarne la somma alla cella della matrice convoluta situata alle coordinate della cella centrale, detta cella focale, della matrice di convoluzione. Scorrendo lungo la matrice convoluta, la cella focale assumerà le coordinate di tutte le celle della matrice convoluta.

Un inconveniente dell'uso dei filtri di convoluzione è che lungo i bordi della matrice più grande (la matrice “convoluta”) non è possibile eseguire il calcolo a causa della assenza di celle contigue la cella focale. Si generano in tal modo dei bordi di valori nulli di dimensioni pari alla metà dell'ampiezza della matrice di convoluzione. Questo è in genere valido anche per le tecniche di neighborhood analysis e tutte le tecniche di analisi *raster* basate su finestre mobili (a meno che non si facciano scorrere per “blocchi” invece che per passi di una cella, ma in tal caso il risultato dell'elaborazione differisce notevolmente)

$$\begin{matrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{bmatrix} & \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 1 & 0 \end{bmatrix} & \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} & \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \\ 0^\circ & 90^\circ & 45^\circ & 135^\circ \end{matrix}$$

Fig 4.2 - Esempio di filtri gradiente (filtri di Prewitt) direzionali

### 4.2.1 Filtri gradiente

Alcuni filtri di convoluzione sono utili per l'individuazione di linee all'interno di immagini *raster*. Le linee, dal punto di vista numerico corrispondono bruschi cambi di intensità nei valori dei pixel. Al fine di identificare tali variazioni si utilizzano filtri che approssimano il valore della derivata prima della funzione continua di valori di intensità luminosa dell'immagine (fig 4.2). La derivata di una funzione continua lungo una determinata direzione, ad esempio lungo la direzione dell'asse  $x$ , risulta la seguente:

$$\frac{\partial F(x, y)}{\partial x} = \lim_{\Delta x \rightarrow 0} \frac{F(x + \Delta x, y) - F(x, y)}{\Delta x}$$

Il limite principale nel voler calcolare la derivata di una funzione da un dato *raster* è costituita dal fatto che esso non costituisce affatto una funzione continua nello spazio ma è una matrice di valori regolarmente spaziatati nel piano  $x$ - $y$ . Si tratta quindi di una versione discreta della funzione di intensità luminosa, per cui non è possibile definire il limite di  $\Delta x$  che tende a zero. Il valore minimo di  $\Delta x$  è pari alla spaziatura dei pixel, che equivale a porre tale limite ad 1 se si utilizza il pixel come unità di misura. Ponendo quindi  $\Delta x=1$ , l'approssimazione della derivata lungo la direzione orizzontale può essere formulata nella maniera seguente

$$\frac{\partial F(x, y)}{\partial x} \approx f(i, j+1) - f(i, j) \quad [\text{con } x=j \text{ e } y=i]$$

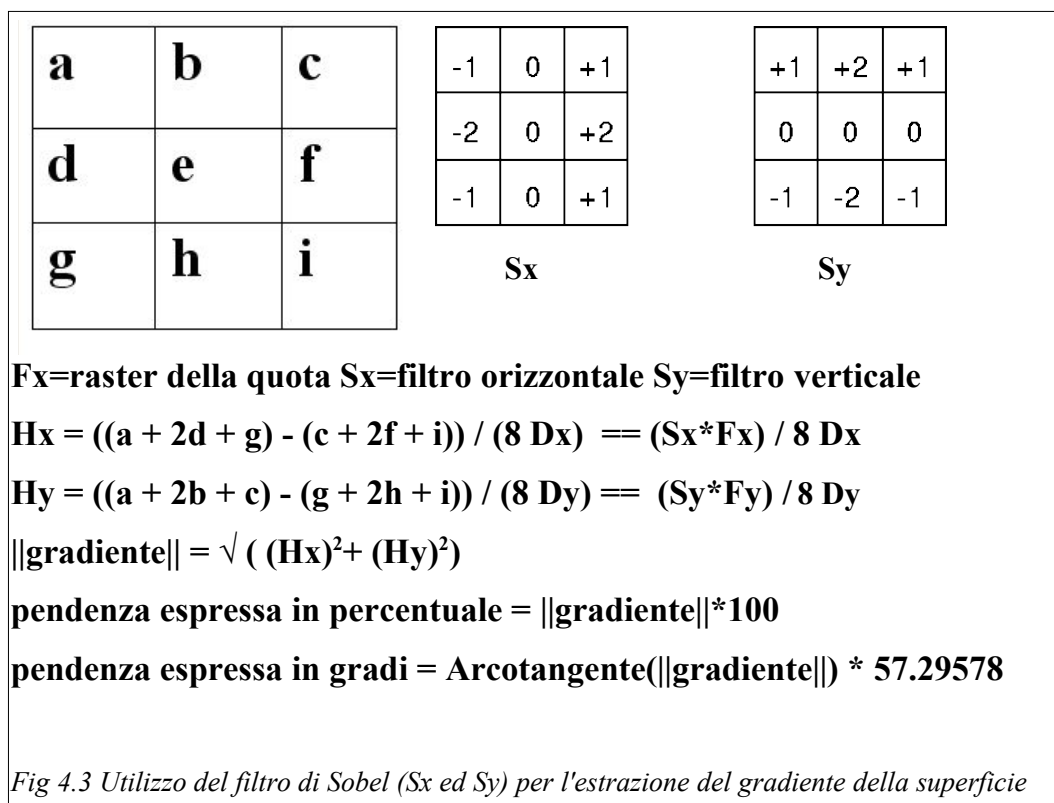
dove  $f(i,j)$  corrisponde al valore del pixel alla "colonna"  $j$  e riga " $i$ " della matrice dei valori.

Naturalmente il sistema di coordinate dipende dal tipo di dato che si vuole analizzare. Nel caso dell'analisi dei dati non geografici è possibile utilizzare un

sistema di coordinate in cui  $x=j$  e  $y=i$ , ed è conveniente considerare il nostro sistema di coordinate  $x,y$  coincidente con la disposizione spaziale di righe e colonne della matrice. Se si analizzano dati spaziali, il nostro sistema di coordinate dovrà coincidere con il sistema di coordinate spaziali. In tal caso la spaziatura delle righe e delle colonne saranno espresse in termini di unità di misura differenti (es. metri, miglia ecc..).

Risulta possibile, volendo utilizzare un'approssimazione migliore della derivata considerare un valore centrale utilizzando una maschera di ampiezza pari a 3 pixel.

$$\frac{\partial F(x, y)}{\partial x} \approx \frac{f(i, j+1) - f(i, j-1)}{2} \quad [\text{con } x=j \text{ e } y=i]$$



Tra i filtri più comunemente utilizzati per il calcolo del gradiente vi sono quelli di Prewitt e di Sobel, che calcolano il gradiente lungo una direzione specifica. Il

gradiente, come specificato nel capitolo 3, è un vettore di L elementi nello spazio 1-dimensionale. Nel caso delle funzioni di due variabili reali è un vettore che punta nella direzione di massima pendenza della superficie ed il cui modulo è appunto la pendenza massima calcolata in quel determinato punto. In fig 4.3 il filtro di Sobel è visualizzato nelle sue due componenti verticale ed orizzontale e le componenti della matrice indicizzate con lettere. Hx e Hy rappresentano le variazioni della funzione nello spazio, ovvero le derivate approssimate dal filtro, e Dx e Dy sono la spaziatura del pixel della matrice. I filtri qui introdotti, sono spesso utilizzati dagli algoritmi per il calcolo dei parametri morfometrici della superficie implementati nei pacchetti Gis. Ad esempio il Sobel definisce uno schema equivalente ai metodi di calcolo proposti da Horn (1981) e Skidmore (1989) per le pendenze e l'aspetto.

Una volta calcolato il gradiente, tramite la formula di fig 4.3, risulta infatti possibile calcolare la pendenza ed aspetto, mentre, per calcolare le derivate di secondo ordine tramite filtri di convoluzione, è necessario applicare nuovamente lo stesso filtro alle matrici che rappresentano Hx e Hy ottenendo le derivate seconde Hxx , Hxy, Hyy. Da queste risulta quindi possibile derivare altri parametri morfometrici descritti nel paragrafo 3. Grazie alla proprietà associativa della convoluzione, è possibile realizzare i filtri che calcolano le derivate seconde senza dover passare due volte lo stesso filtro sul DTM.

Per esempio, utilizzando i filtri di Sobel per ottenere Hxx si può realizzare il filtro  $S_{xx}=S_x*S_x$  ed applicarlo ad Fx, che equivale a “passare” due volte il filtro  $S_x$  sul *raster* Fx;  $H_{xx}=(S_x*S_x)*F_x=S_x*(S_x*F_x)$  similmente  $H_{xy}=(S_x*S_y)*F_x$ .

Una volta definiti i filtri direzionali, questi vengono applicati secondo uno schema a finestre mobili con passo di un pixel:

$$H_x(i, j) = 1/(D_x * 8) \sum_{a=A_{start}}^{a=A_{end}} \sum_{b=B_{start}}^{b=B_{end}} S_x(a, b) f(i-a, j-b)$$

$$H_y(i, j) = 1/(D_y * 8) \sum_{a=A_{start}}^{a=A_{end}} \sum_{b=B_{start}}^{b=B_{end}} S_y(a, b) f(i-a, j-b)$$

Dove  $S_x$  e  $S_y$  sono i filtri 3x3 di Sobel di fig 4.3 rispettivamente orizzontale e verticale, e  $D_x$  e  $D_y$  sono la spaziatura orizzontale e verticale del *raster* espressa in metri. Le differenze  $b_{end}-b_{start}$  e  $a_{end}-a_{start}$  sono entrambi pari alla dimensione della maschera, ovvero 3. In questa notazione è presente come divisore il fattore di scala, ovvero  $D_x$  e  $D_y$ , ed il fattore di normalizzazione del prodotto della convoluzione, essendo  $[1+2+1] \times 2 = 8$ .

#### 4.2.2 Filtri di smoothing

Una delle tecniche maggiormente utilizzate per ridurre gli artefatti presenti nei DTM è lo *smoothing*. Si tratta di un'operazione mirata a ridurre il rumore di fondo presente nelle immagini *raster*. Per rumore di fondo si intende quella componente delle frequenze alte che non è dovuta a proprietà del microrilievo, ma a difetti di campionamento o di interpolazione. Le componenti alte e basse presenti in un DTM possono essere analizzate tramite trasformata di Fourier. La trasformata discreta di Fourier (DFT) permette di trasformare il segnale dal dominio spaziale in dominio delle frequenze. In tale dominio è possibile visualizzare lo spettro delle frequenze che corrisponde ad un grafico per un segnale 1D ed un *raster* per un segnale 2D. Attraverso l'analisi dello spettro dei filtri di convoluzione è possibile identificare quali frequenze verranno tagliate nel *raster* convoluto. Infatti la convoluzione tra due segnali nel dominio spaziale equivale ad una moltiplicazione nel dominio delle frequenze. Si possono quindi definire filtri passa-alto che tagliano le frequenze più basse, filtri passa-basso che tagliano le frequenze più elevate e filtri passa-banda che tagliano specifiche fasce di frequenze.

Il più semplice dei filtri di smoothing è il cosiddetto Pillwitt, ovvero un filtro che utilizza una matrice con valori pari ad 1 ed un fattore di divisione pari al numero di celle del filtro.

Altri filtri più sofisticati hanno simmetria rotazionale e sono definiti tramite rotazione di una funzione unidimensionale  $\gamma(r)$  dove  $r = \sqrt{a^2+b^2}$  e costituisce la

distanza tra il centro del *kernel* e gli elementi della maschera a e b. Un tipico filtro a simmetria rotazionale è il filtro gaussiano.

La funzione gaussiana è definita come

$$y(r) = \frac{1}{2\pi\sigma^2} e^{\frac{-1}{2} \left(\frac{r}{\sigma}\right)^2}$$

mentre la corrispondente versione discreta (in forma matriciale) è definita come

$$g(a, b) = \frac{1}{2\pi\sigma^2} e^{\frac{-1}{2} \frac{a^2 + b^2}{\sigma^2}}$$

Le variabili a e b sono considerate variabili reali e quindi definite nell'intero piano, anche se in realtà il filtro opera solo su di un intervallo definito poiché gli elementi, a ed b, sono celle del *raster* (rispettivamente la riga a-esima e la colonna b-esima della matrice di convoluzione).

Il fattore di normalizzazione  $1/(2\pi\sigma^2)$  fa sì che l'integrale della funzione sia uguale ad uno.

Il valore di deviazione standard è espresso in numero di pixel del filtro di convoluzione, risulta quindi indispensabile utilizzare un numero  $\sigma$  maggiore o uguale a 1 essendo altrimenti non efficace.

Grazie al fatto che i pixel più prossimi al centro della maschera assumono maggior peso rispetto ai pixel più lontani, tale filtro riesce a produrre uno *smoothing* di qualità superiore rispetto ad esempio ad un semplice filtro mediano. Essendo una funzione irrotazionale è evidente che uno *smoothing* gaussiano risulta meno influenzato dalla struttura a griglia regolare rispetto ad altri tipi di filtro che non tengono conto della diversa distanza dei pixel diagonali.

2	4	5	4	2
4	9	12	9	4
5	12	15	12	5
4	9	12	9	4
2	4	5	4	2

Fig 4.4 Versione discreta del filtro gaussiano con  $\sigma=1.0$

### 4.2.3 Filtro laplaciano

Un'altro filtro importante è il laplaciano del gaussiano. Il laplaciano è la funzione definita come

$$L(x, y) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

Nel caso delle immagini digitali è utilizzato per identificare rapidi cambiamenti di intensità luminosa. Essendo una funzione dipendente dalle derivate seconde viene utilizzato in congiunzione con uno *smoothing* gaussiano al fine di ridurre l'effetto di sensibilità al rumore di fondo. Il filtro risultante dalla convoluzione di un filtro laplaciano ed un filtro gaussiano è detto appunto laplaciano del gaussiano o LoG in inglese. Poiché la convoluzione è un'operazione associativa risulta più semplice effettuare la convoluzione tra le maschere più piccole e poi applicare il filtro risultante al *raster* più grande.

0	1	1	2	2	2	1	1	0
1	2	4	5	5	5	4	2	1
1	4	5	3	0	3	5	4	1
2	5	3	-12	-24	-12	3	5	2
2	5	0	-24	-40	-24	0	5	2
2	5	3	-12	-24	-12	3	5	2
1	4	5	3	0	3	5	4	1
1	2	4	5	5	5	4	2	1
0	1	1	2	2	2	1	1	0

Fig 4.5 Versione discreta del LoG con  $\sigma=1.4$



Questo filtro è praticamente in grado di evidenziare i cambiamenti di curvatura media, da concava a convessa di un DTM. Infatti, come visto nel capitolo 3, la media della somma di due curvature calcolate lungo due sezioni mutuamente perpendicolari corrisponde alla curvatura media in quel determinato punto. Dalla formula della curvatura media  $H$  (Shary et alii 2002) si evince che per superfici caratterizzate da basse pendenze, il valore risultante dall'approssimazione ricavata dal LoG, tende a due volte la curvatura media. Difatti (utilizzando le stesse notazioni del paragrafo 3.5) se  $p$  e  $q$  tendono a zero,  $r+t$  che è appunto il laplaciano, tende a  $-2H$ . Aumentando il valore di sigma del filtro gaussiano è possibile ottenere un maggior grado di tolleranza, poiché le brusche variazioni di quota vengono attenuate dallo *smoothing* gaussiano. Tramite il filtro LoG si può inoltre realizzare un'analisi multiscala delle concavità e convessità medie modificando il raggio di estensione del filtro gaussiano.

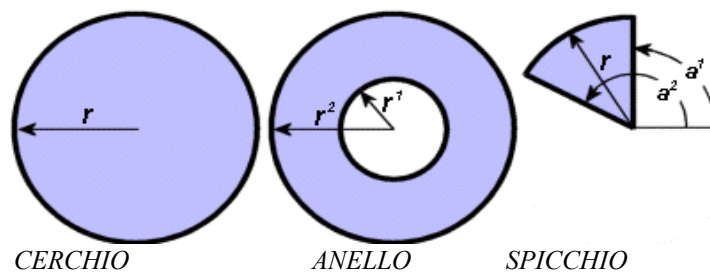
#### **4.4 Neighborhood Analysis**

Il metodo della convoluzione è solo una delle tecniche possibili di *neighborhood analysis*. Altri metodi di analisi di vicinanza prevedono l'uso di operazioni non basati sul concetto di “point spread function” ed anche l'uso di finestre mobili dalla forma differente dalla quadrata. Inoltre possono essere fatte scorrere secondo uno schema differente rispetto al metodo di convoluzione, ad esempio per blocchi adiacenti o sovrapposti ma per passi maggiori di un pixel. Filtri utilizzati nell'analisi di vicinanza sono per esempio quelli di tipo statistico (Tab. 4.1). Anche in questi filtri il valore ottenuto dall'operazione è associato ad una cella focale la cui posizione dipende dalla forma del filtro utilizzato.

## Metodi statistici in arcview

- Majority: determina il valore più frequente
- Minority: determina il valore meno frequente
- Maximum: valore massimo
- Minimum: valore minimo
- Range: differenza tra valore massimo e valore minimo
- Mean: totale dei valori diviso numero di celle.
- Median: valore mediano
- Standard Deviation: calcola la dispersione del dato
- Sum: calcola il valore sommando il totale delle celle
- Variety: determina il numero di valori differenti

*Tabella 4.1 Analisi statistiche comunemente disponibili tramite la "neighborhood analysis" (da ArcView Spatial Analyst Manual (c) Esri 1996)*



*Fig 4.6. Vari tipi di forme di filtri neighborhood (da ArcView Spatial Analyst Manual (c) Esri 1996)*

### 4.5 Lo shaded relief

Lo “shaded relief” o “Hill shading” (Horn 1981) è una tecnica di visualizzazione che permette di percepire in maniera efficace la forma della superficie senza la necessità di impiegare metodi di visualizzazione tridimensionale. Questo tipo di visualizzazione viene spesso indicato come “analytical hillshading” poiché risulta molto utile ad evidenziare strutture regolari o meno nell'area di studio. Spesso tale tecnica di visualizzazione permette di evidenziare molto bene forme di origine strutturale ed quindi è stata impiegata

come ausilio allo studio dell'evoluzione tettonica (Jordan 2003, Ganas et alii 2005). Gli algoritmi utilizzano una fonte luminosa definita da vari parametri. Poiché la fonte luminosa è considerata essere posta a distanza infinita, l'algoritmo è molto semplificato rispetto ai metodi di illuminazione impiegati negli algoritmi di *ray tracing* (Newmann & Sproull 1979), che permettono un tracciamento più accurato delle ombre. Tale metodo risulta disponibile come opzione in alcuni pacchetti specialistici, cioè non veri e propri GIS, quali ad esempio DIGEM (Conrad 2000).

Le variabili utilizzate sono l'azimut della fonte luminosa, che può ovviamente essere modificato a piacere rispetto alla reale posizione geografica dell'area di studio, la declinazione, che definisce l'angolo di incidenza sulla superficie, la pendenza del DTM, che per limiti dell'algoritmo non può superare i 90 gradi, ed ovviamente l'aspetto della superficie del DTM che varia da 0 a 360 gradi (si utilizza la solita convenzione di identificare il nord con 90 ed il sud con 270). L'intensità luminosa del singolo pixel del DTM è calcolata tramite la seguente formula (Eyton 2005):

$$R_f = \cos(\alpha_f - \alpha_s) \sin(Y_f) \sin(Y_s) + \cos(Y_f) \cos(Y_s)$$

Dove  $R_f$  è la riflettanza della cella  $\alpha_f$  ed  $\alpha_s$  sono gli angoli che definiscono la direzione dell'aspetto del DTM e della fonte luminosa ed  $Y_f$  e  $Y_s$  sono rispettivamente la pendenza della superficie e l'angolo di declinazione del sole. Il valore varia tra 0 e 255. Nel caso il risultato sia di segno negativo viene posto a zero, in maniera tale da simulare la presenza di un'ombra. Ovviamente l'ombra non è reale poiché non tiene conto della quota delle celle circostanti che possono bloccare il percorso del raggio luminoso ed oscurare aree più vaste, calcolo invece previsto dagli algoritmi di *ray tracing*, ma in tal modo si verifica il vantaggio di poter orientare a piacimento la fonte luminosa in maniera da individuare, tramite l'allineamento delle ombre anche i lineamenti e quindi le forme a controllo strutturale, come ad esempio i versanti di faglia.

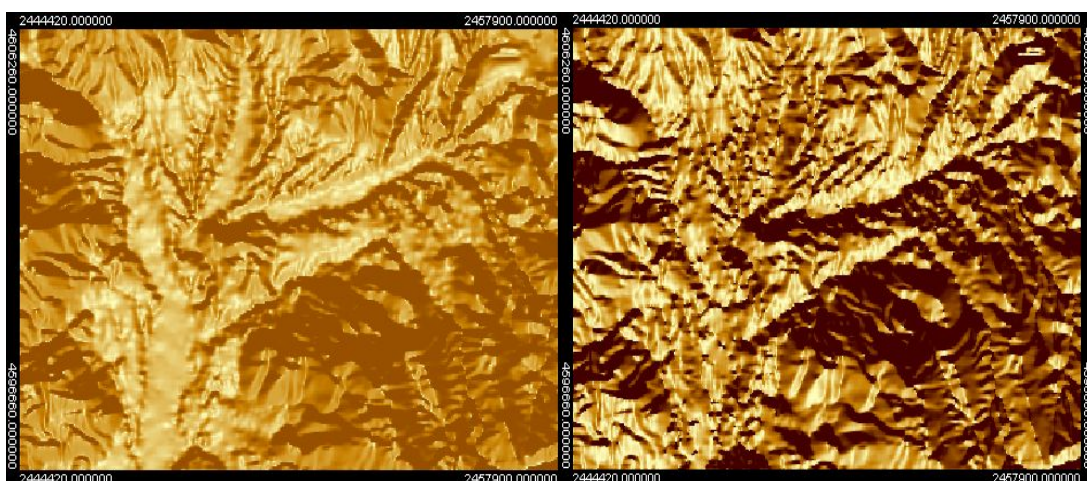


Fig 4.7 differenza tra raster dell'aspetto (a destra) con uno shaded relief (a sinistra) realizzati tramite il software DIGEM. Notare la differenza nella resa visiva delle aree a bassa pendenza dovuta al calcolo della riflettanza.

#### 4.6 Le operazioni locali e la map algebra.

Le operazioni locali consistono un'insieme di tecniche per trasformare il valore di singole celle .

Si utilizzano un o più *raster* per ottenere un nuovo *raster* ove i valori delle singole celle sono il risultato di una operazione logica od algebrica. È possibile così effettuare trasformazioni lineari o non lineari utilizzando qualsiasi tipo di funzione applicata alla cella del *raster* o dei *raster* di partenza ed assegnando il valore alla cella del *raster* risultante. Lo strumento principale per effettuare tali operazioni viene chiamato *map-algebra* che può essere dotato di vari operatori logici, aritmetici, trigonometrici ecc.. ed in alcuni casi utilizzabile in maniera iterativa tramite appositi script. Ogni software GIS ha i suoi strumenti di *map*

*algebra*, tuttavia tra i più flessibili vi è quello implementato nel GIS utilizzato per gli elaborati di questa tesi (Shapiro & Westervelt 1992) la cui sintassi è molto semplice ma allo stesso tempo estremamente flessibile

*r.mapcalc risultato=espressione*

Un esempio banale di utilizzo della map algebra può essere ad esempio una trasformazione non lineare dell'intero *raster*. Si abbia ad esempio il caso in cui i valori della mappa che si sta elaborando sono caratterizzati un'estensione numerica che, associata ad una scala cromatica lineare, non permette di percepire visivamente in maniera adeguata l'andamento della variabile, od, altro esempio, il caso di una distribuzione non gaussiana di una variabile che si vuole impiegare in analisi di tipo multivariato che hanno come presupposto una relazione gaussiana tra le variabili. Prendendo il caso della gamma di intervalli lineare, un'operazione locale che effettua una trasformazione logaritmica rende possibile evidenziare nella scala cromatica anche i valori estremi:

*r.mapcalc nuovo\_raster=log(raster1)*

se alcuni valori del *raster* di partenza sono inferiori allo zero il logaritmo non è definito. L'utente vorrebbe ad esempio poter trasformare comunque il *raster* e conservare il segno della variabile. In tal caso potrebbe o “shiftare” i valori oltre lo zero utilizzando altre funzioni e sfruttando la flessibilità della map algebra:

*r.mapcalc nuovo\_raster=(abs(raster1)/raster1)\*log(abs(raster1))*

Altre operazioni realizzabili tramite map-algebra è il *map overlay* ovvero operazioni che utilizzano due o più *raster*, e sono impiegate in ambiente GIS in

numerosissimi campi delle scienze naturali, ingegneria, urbanistica ecc. Benché esistano appositi comandi all'interno dei pacchetti GIS (ad esempio *r.combine* ed *r.weight* nel *gis GRASS*), la *map algebra* risulta molto flessibile in tal senso. Può essere utilizzata per realizzare combinazioni lineari e non di numerosi *raster*. Un esempio banale è il calcolo del fattore del modello *RUSLE* (Mitas et alii 1996a):

*r.mapcalc A=R\*K\*L\*S\*C\*P*

dove le lettere indicano i *raster* che contengono i valori dei fattori *RUSLE*.

#### **4.7 Le operazioni zonali**

Mentre l'analisi di vicinanza esegue operazioni basate sul concetto di cella focale, esiste un tipo di analisi per insiemi di celle del *raster* basata invece che dalla vicinanza spaziale, dal condividere lo stesso attributo. Tali insiemi di celle vengono definite “zone”. Quando si effettua un'analisi statistica zonale, si utilizza un *raster* cosiddetto di “copertura”, ossia un *raster* che rappresenta una caratteristica continua della regione del *raster*, come ad esempio la quota o la pendenza, dal quale estrarre i valori.

Per le celle caratterizzate dallo stesso attributo numerico o categorico viene calcolato il valore statistico, utilizzando le celle del *raster* di copertura che ricadono entro ogni zona. Si genera così un nuovo *raster* e/o una tabella, che associa ad ogni zona il valore statistico ottenuto.

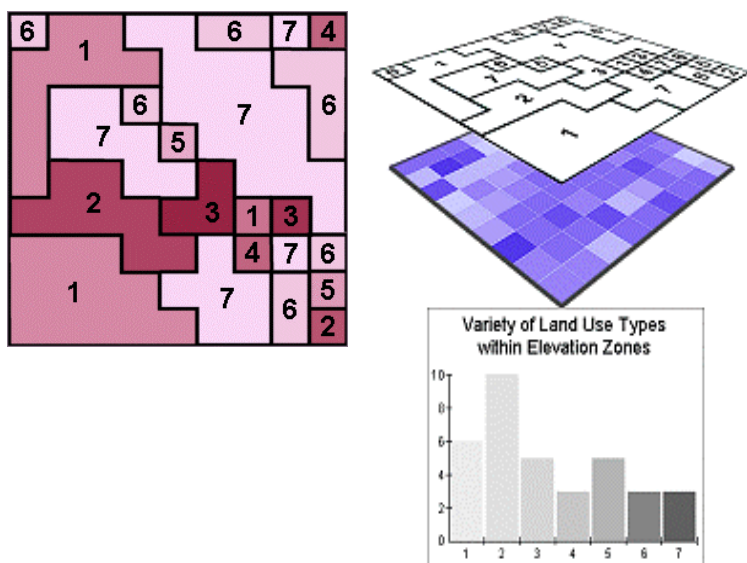


Fig 4.8 Esempio di analisi zonale (da ArcView spatial analyst manual, 1996)

## B.III Classificazione del pattern

### 5.1 Introduzione

In questo paragrafo si introducono i concetti di pattern e classificazione del pattern. Il problema della classificazione, qualsiasi sia il campo di applicazione o l'oggetto della classificazione, è scomponibile in un diagramma di flusso di (fig. 5.1).

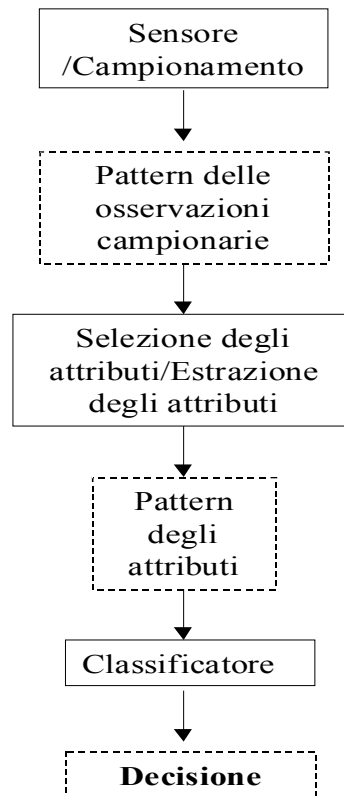


Fig. 5.1 - Fasi del processo di classificazione

(da Webb 2002 ridisegnato e modificato)

Lo schema è generico e le varie fasi illustrate possono a seconda del problema da affrontare o della metodologia impiegata, essere assenti o inglobate in un unico passaggio. Comunque prima di poter decidere quali siano le fasi e come ogni fase del processo debba essere realizzata è indispensabile definire la natura del problema. Nel caso del presente lavoro, la natura del problema è stata già definita nel capitolo introduttivo. Ciò consente di poter scegliere come



realizzare le successive fasi:

-Raccolta/campionamento del dato reale: è la fase che consente, tramite strumentazione appropriata, di campionare le variabili del fenomeno naturale che nel loro insieme andranno a costituire la struttura o pattern che si vuole classificare. Si deve sottolineare che la terminologia in questo caso può risultare ambigua poiché si identifica come pattern l'insieme dei dati campionati, o in genere qualsiasi “struttura” di un insieme di dati, mentre dal punto di vista matematico per pattern esiste in letteratura una definizione specifica. Ciò non ha rilevanza dal punto di vista della strutturazione delle fasi essendo solo un problema di terminologia. L'importante è capire la differenza tra un procedimento e l'altro.

-Analizzare in maniera preliminare il *pattern*: ad esempio estrarre le statistiche di base (media, varianza ecc..), visualizzare diagrammi quali gli istogrammi della distribuzione, *scatterplot* od altri tipi di visualizzazione bi e tridimensionale ecc.. al fine di verificare la qualità e come il dato è strutturato. Questa fase è molto importante al fine della scelta delle variabili e del classificatore da utilizzare.

-Selezione e/o estrazione degli attributi: la selezione e l'estrazione sono due procedure differenti entrambi mirati alla riduzione del numero di variabili e quindi della dimensionalità del problema. La scelta consente di utilizzare solo le variabili più utili alla classificazione mentre l'estrazione è un processo che consente di produrre, a partire da un certo numero di variabili, tramite trasformazioni lineari o non lineari, un nuovo set di variabili. In genere le nuove variabili estratte sono di dimensionalità ridotta rispetto a quelle originali, in maniera da ottenere anche in questo caso una riduzione del dato.

-Classificazione: la scelta del classificatore è cruciale poiché esistono moltissime tecniche di classificazione del pattern. Si distinguono due tipologie principali, ovvero la classificazione supervisionata e quella non supervisionata.

*La classificazione non supervisionata, o clustering, permette di effettuare analisi esplorative del set di dati in maniera da poter evidenziare strutture interessanti utili alla realizzazione di un classificatore specifico per il problema. Spesso il risultato del clustering, in combinazione con una opportuna fase di scelta ed estrazione di attributi, risulta già soddisfacente quale metodo di classificazione. Altre volte è necessario utilizzare metodi di classificazione supervisionata, ovvero utilizzare un sottoinsieme del set di dati scelti dall'operatore da utilizzare come “training dataset”. In entrambi i tipi di classificazione è comunque determinata la scelta della metodologia da utilizzare visto che a partire dallo stesso insieme di attributi e “training dataset” si possono ottenere risultati molto differenti.*

-Validare il risultato ottenuto: consente di verificare nel caso si utilizzi un classificatore di tipo supervisionato, la corrispondenza tra classi individuate manualmente (e quindi note a priori) e classi ottenute. Risulta invece necessario interpretare il risultato della classificazione non supervisionata, poiché come vedremo, a scapito di una maggiore oggettività del risultato ottenuto, i metodi di *clustering* si limitano ad identificare strutture del set di dati che devono poi essere interpretate dall'operatore. La validazione in questo caso può anche limitarsi ad una valutazione della interpretazione del risultato.

## **5.2 Definizione di pattern**

Dal punto di vista formale il *pattern* (Webb 2002) consiste in una osservazione all'interno di un set di dati di cui sono state misurate un numero  $d$  di variabili (detti anche attributi o *features*). Il pattern può essere rappresentato come un vettore  $\mathbf{x}$  di  $d$  elementi. Il vettore  $\mathbf{x}_i$  costituisce quindi un vettore di attributi dell'osservazione  $i$ -esima. Il numero di attributi  $d$ , definisce la dimensionalità dello spazio degli attributi (o dimensionalità del *pattern*). Un

insieme di pattern  $L$  consiste nell'insieme dei vettori degli attributi  $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$  mentre il *pattern*  $i$ -esimo è il vettore  $\mathbf{x}_i = [x_{i,1}, \dots, x_{i,d}]^T$  ove  $T$  denota trasposizione. In tal modo l'insieme dei pattern può essere rappresentato tramite una matrice  $\mathbf{X}$ .

Tutte le variabili estratte dai DTM costituiscono dati spaziali, quindi i rispettivi pattern consistono in osservazioni originariamente georeferenziate. Data la struttura a griglia regolare del *raster*, tali osservazioni sono costituite dagli attributi associati alle celle del *raster*. In questo caso  $n$ , il numero di osservazioni, corrisponde al numero di celle del DTM, mentre  $d$  corrisponde al numero degli attributi associati ad ogni cella del *raster*. Il set di dati da elaborare consiste, quindi, in una matrice di dimensioni  $n \times d$ . In tal caso l'informazione spaziale (georeferenziazione) non è inclusa nello spazio dei vettori delle variabili. Le coordinate geografiche non vengono quindi utilizzate nel processo di classificazione.

Difatti, le librerie comunemente disponibili per l'analisi multivariata, utilizzano algoritmi che non tengono conto della collocazione nello spazio geografico, ma solo nello spazio degli attributi. Tale problematica è stata affrontata mediante il metodo descritto nella parte C.

### 5.3 Tecniche di classificazione supervisionata

Esistono in letteratura numerose tecniche di classificazione supervisionata. Nella maggior parte dei casi la classificazione supervisionata è stata applicata a immagini multispettrali telerilevate da sistemi satellitari (es. Landsat e SPOT), costituite da matrici di pixel. Ciascun pixel è caratterizzato da una distinta “firma spettrale”, vale a dire da un vettore di valori di radianza o riflettanza acquisiti a diverse lunghezze d'onda. La classificazione supervisionata di un'immagine multispettrale richiede la delimitazione di aree test (o *training areas* o *region of interest*) spettralmente omogenee, rappresentate cioè da gruppi di pixel con strette analogie spettrali (statisticamente verificate). In tal caso, il concetto di “firma

spettrale”, già applicato al singolo pixel, viene esteso all'intera area test.

Per similitudine con il concetto di “firma spettrale” una sorta si può definire di “firma geometrica” della superficie. Tale concetto è stato introdotto da Pike (1988) e costituisce un insieme di misure che consentono di descrivere la superficie topografica in maniera sufficientemente dettagliata da poter distinguere i vari elementi del paesaggio. Brown et alii (1998) hanno sperimentato e confrontato tra loro l'efficacia della classificazione basata sulla massima verosimiglianza (Maximum Likelihood o ML) e quella basata sulle reti neurali artificiali (ANN). Esistono numerosi altri tipi e varianti di classificatori di tipo supervisionato e non supervisionato. Per un'analisi completa delle caratteristiche dei vari tipi di classificatori si rimanda alla bibliografia. La più semplice tecnica di ML, ed utilizzata nel lavoro menzionato, si basa sulla probabilità di appartenenza del vettore degli attributi (feature vector) ad una determinata classe. La semplicità del metodo si basa sul fatto di considerare note a priori sia il tipo di distribuzione, che i parametri della stessa. Nel caso della ML le classi sono caratterizzate da una distribuzione gaussiana ed i parametri sono noti a priori perché estratti dal “training dataset” cioè dai vettori degli attributi associati ai pixel selezionati dall'utente. Quindi attribuire un pixel ad una determinata classe significa semplicemente calcolare le matrici di covarianza e la media delle classi, ed assegnare ogni vettore degli attributi alla classe per la quale risulta massima la probabilità condizionata secondo la predetta distribuzione gaussiana.

Benché il metodo della ML sia ottimale dal punto di vista probabilistico, il suo utilizzo non risulta così immediato. Il classificatore deve attribuire ogni pixel ad una e soltanto una classe. Questo in base al criterio di Bayes. Precisamente tale criterio consiste nell'assegnare un pattern alla classe che ha la probabilità condizionata maggiore rispetto al vettore. Dato quindi un vettore degli attributi  $\mathbf{x}$ , avremo che  $\mathbf{x}$  è un elemento appartenente alla classe  $w_i$  se e soltanto se

$P(w_i|\mathbf{x}) > P(w_k|\mathbf{x})$  per  $k=1, \dots, C$  e  $k \neq i$  dove  $C$  è il numero di classi esistenti

In pratica non esiste la possibilità che un pixel non appartenga a nessuna delle

classi identificate dall'operatore. Questo si riscontra difatti nei software che utilizzano tale metodo. Per cui per poter classificare in maniera corretta un *raster* è necessario identificare manualmente un “training dataset” per tutte classi presenti nell'area da classificare.

Altri classificatori più semplici sono basati sulla distanza del vettore degli attributi dalle classi. In questo caso un vettore è attribuito alla classe dalla quale dista meno.

Le due distanze più comunemente utilizzate sono la distanza euclidea definita come

$$De^2 = \|x_i - \mu_i\|^2$$

che corrisponde alla distanza tra vettore e classe, per classi con uguale matrice di covarianza e matrice diagonale. Nel caso più frequente di classi con matrici di covarianza non diagonali, si utilizza la distanza di Mahalanobis

$$Dm^2 = (x_i - \mu_i)^T \Sigma^{-1} (x_i - \mu_i)$$

Nella parte C si utilizzerà il metodo della minima distanza statistica per individuare alcune forme del rilievo in maniera supervisionata, un metodo più semplice rispetto alla ML, ma comporta dei vantaggi rispetto a quest'ultima, come descritto nel paragrafo 7.3, almeno nell'ambito della cartografazione automatica delle forme del rilievo ove non è noto a priori il numero di classi presenti.

## 5.4 Tecniche di Classificazione non supervisionata

Per classificazione non supervisionata o *clustering* si intendono delle tecniche in grado di suddividere un set di dati in un certo numero di classi senza l'intervento diretto dell'operatore durante la fase di riconoscimento. L'unico intervento consiste nel definire alcuni parametri prima della classificazione automatica.

Nel *clustering*, ogni osservazione proveniente dal set di dati ha le proprie

coordinate in uno spazio multivariato, determinato dall'insieme delle variabili del set di dati. I punti che si trovano "vicini" in questo spazio vengono raggruppati in un unico *cluster*. Le tecniche per determinare i *cluster* sono varie, in particolare si differenziano le tecniche di tipo gerarchico da quelle di tipo non gerarchico o partizionale.

### ***Clustering partizionale***

La tecnica di *clustering* partizionale più diffusa si basa sulla minimizzazione del costo di una determinata funzione. I più popolari algoritmi di *clustering* iterativo sono il K-means ed il fuzzy c-means (Webb 2002). I parametri della funzione sono i prototipi dei *cluster* - ossia un parametro che serve a rappresentare un determinato *cluster*- la matrice delle distanze tra i vettori ed i prototipi dei *cluster* ed un criterio,  $d(x_i, \Theta_j)$ , che definisce la distanza o dissimilarità tra i vettori degli attributi ed i prototipi dei *cluster*. In base al tipo di distanza multivariata ed ai prototipi utilizzati si possono individuare *cluster* dalla forma differente.

L'obiettivo del fuzzy c-means è quello di minimizzare la seguente funzione costo

$$J_q(\Theta, U) = \sum_{i=1}^N \sum_{j=1}^m u_{ij}^q d(x_i - \Theta_j)$$

con

$\Theta_j$  = vettore prototipo del cluster  $J$ -esimo

$U$  = matrice dei vettori  $N \times m$  i cui elementi  $(u_{ij})$  sono rispettivamente

$u_j(x_i)$  ed  $d(x_i, \Theta_j)$

mentre  $q$  è un parametro detto fuzzifier

nel fuzzy c-means il valore membro  $u_{ij}$  è soggetto ai vincoli

$$\sum_{j=1}^m u_{ij} = 1 \quad \text{con} \quad i = 1, \dots, N$$

ed

$$u_{ij} \in [0, 1], \quad i = 1, \dots, N, \quad j = 1, \dots, m; \quad 0 < \sum_{i=1}^N u_{ij} < N, \quad j = 1, 2, \dots, m$$

In pratica il valore membro  $u_{ij}$  definisce il “grado” di appartenenza del vettore  $i$ -esimo al *cluster*  $j$ -esimo ed è vincolato a tutti gli altri  $j-1$  valori membro, cioè al grado di appartenenza al resto dei *cluster*. La somma dei valori membro per ogni vettore è pari ad 1. In pratica il vettore  $i$ -esimo viene ad essere associato contemporaneamente a tutti i  $J$  *cluster* con un grado di appartenenza che varia tra zero ed uno. Un valore pari a zero significa che quel vettore non ha nessuna affinità con quel *cluster* mentre un valore pari ad uno ha il significato molto preciso, dato il vincolo per cui i valori membro devono sommare ad uno, di appartenere unicamente a quel *cluster* dato che i restanti valori membro dovranno essere evidentemente pari a zero.

Per definire la funzione obiettivo è necessario definire prototipo del *cluster*, ossia con quali parametri viene rappresentato il *cluster* nello spazio multivariato. Il metodo del fuzzy c-means si basa sulla rappresentatività dei *cluster* in base ai centroidi degli stessi. Questi sono definiti come

$$\Theta_j = \frac{\sum_{i=1}^N u_{ij}^q \mathbf{x}_i}{\sum_{i=1}^N u_{ij}^q}$$

mentre si utilizza come criterio di dissimilarità,  $d(\mathbf{x}_i, \Theta_j)$ , la distanza euclidea tra il vettore ed il centroide definito come sopra. In tal modo l'algoritmo fuzzy c-means, una volta definiti il numero di *cluster*  $m$  ed il valore di “fuzzifier”, può essere descritto come di seguito

1 – si inizializzano i centroidi degli  $m$  *cluster*  $\Theta_j \quad J=1....m$  in base ad un criterio conveniente (ad esempio si possono scegliere  $m$  vettori a caso dal set di dati ed utilizzare le loro coordinate come seme iniziale)

2 – si pone il numero di iterazioni  $t=0$

3 – si calcola la matrice dei valori membro come

$$u_{ij} = \frac{1}{\sum_{k=1}^m \left( \frac{\|x_i - \Theta_j(t)\|^2}{\|x_i - \Theta_k(t)\|^2} \right)^{\frac{1}{q-1}}}$$

( $\|x_i - \Theta_j(t)\|^2$  denota distanza euclidea tra il vettore  $i$ -esimo ed il centroide del cluster  $j$ )

4- si aggiorna il numero di iterazioni  $t=t+1$

5- per  $j$  che va da 1 ad  $m$  si aggiornano i centroidi

$$\Theta_j = \frac{\sum_{i=1}^N u_{ij}^q(t-1) x_i}{\sum_{i=1}^N u_{ij}^q(t-1)}$$

6 – si ripete dal punto 3 finché  $J(\Theta_j(t), U(t)) - J(\Theta_j(t-1), U(t-1)) < \epsilon$

cioè la differenza di costo tra due iterazioni successive è minore di un determinato valore di tolleranza, oppure l'algoritmo può fermarsi dopo un prestabilito numero di iterazioni.

Una volta che si arresta la procedura è possibile effettuare un'operazione detta di “thresholding”. In pratica si estrae un risultato di tipo “hard”. Tale operazione consiste nell'assegnare il vettore  $i$ -esimo al cluster  $j$ -esimo che presenta il valore membro più grande ovvero  $u_{ij} > u_{ik}$  per  $j=1\dots m, k=1\dots m | j \neq k$ .

Vengono quindi definite delle "Hard Classes" ossia le classi a cui assegnare i pixel. Attraverso la matrice dei valori membro è possibile definire il grado di incertezza (o confusione) nella classificazione dei pixel.

In pratica sia il k-means che il fuzzy c-means si basano sulla minimizzazione dell'errore quadratico e ciò equivale ad utilizzare una distanza euclidea come misura di dissimilarità. La funzione obiettivo del fuzzy k-means è definita come



per il fuzzy c-means con la differenza che il valore membro può assumere solo i due valori di 0 oppure 1. In pratica un vettore può appartenere soltanto ad un *cluster*. I vettori vengono quindi assegnati al *cluster* per il quale la distanza euclidea è minore rispetto a tutti gli altri *cluster*.

$$u_{ij} = \begin{cases} 1 & \text{se } \|x_i - \Theta_j\|^2 = \min \text{ per } j = 1 \dots m \\ 0 & \text{altrimenti} \end{cases}$$

in questo caso ponendo  $\Theta_j$  centroide del *cluster* j-esimo e definendo la misura di distanza come la distanza euclidea tra il vettore ed il centroide è definita come

$$\sqrt{\sum_{l=1}^d (x_l - \Theta_l)^2} \quad \text{con } d = \text{numero di elementi del vettore}$$

il centroide si aggiorna tramite la formula

$$\Theta_j = \frac{\sum_{i=1}^N (t-1) x_i}{N}$$

con N numero di vettori associati al *cluster* J-esimo (cardinalità del *cluster*) all'iterazione (t-1). Anche in questo caso il processo si può arrestare se la differenza tra il costo della funzione tra due iterazioni successive è inferiore ad un valore di tolleranza specifico. Poiché si utilizza la distanza euclidea tale costo equivale alla somma degli scarti quadratici dei singoli *cluster*.

In genere per il K-means il criterio utilizzato per bloccare il processo si basa più frequentemente sulla riallocazione dei vettori tra i vari *cluster* poiché si può escludere dall'algoritmo il calcolo della matrice membro che risulta superflua. Ad ogni iterazione, dopo il calcolo dei nuovi centroidi, si ricalcola la distanza tra vettori e nuovi centroidi. In base a tale distanza, i vettori possono essere assegnati ad un *cluster* diverso oppure non essere riallocati. Il processo si blocca inevitabilmente quando non è più possibile spostare i vettori da un *cluster* all'altro. Tale caratteristica del K-means fa sì che la procedura converga spesso verso minimi locali. Il fuzzy c-means è meno soggetto a tale inconveniente anche se risulta comunque soggetto a problemi di convergenza, il cui limite deve comunque essere definito. Si tratta in entrambi i casi di algoritmi la cui

caratteristica è quella di essere efficienti dal punto di vista dei tempi di calcolo ma il cui risultato si può considerare sub-ottimale. Algoritmi differenti, quali il “simulated annealing” (cristallizzazione simulata, [Kirkpatrick et alii 1983]), cercano di raggiungere un minimo globale ma risultano molto complessi dal punto di vista dei tempi di elaborazione.

Nei lavori di Irvin et alii (1997) e Bourrough et alii (2000), si espone l'uso di algoritmi per il *clustering* iterativo quali ISODATA e il fuzzy c-means al fine di partizionare un DEM in poche classi associabili ad alcune specifiche unità del paesaggio caratterizzate da vari parametri morfometrici. Nell'utilizzare algoritmi di tipo “Hard” come il k-means o la sua variante ISODATA (Ball & Hall 1965) - una versione integrata in GIS quali ArcInfo - non viene però affrontato il problema della classificazione dal punto di vista della sovrapposizione degli attributi nello spazio geografico. Nel caso del fuzzy c-means, la sovrapposizione è gestita solo nello spazio dei vettori degli attributi (vettori dei parametri morfometrici).

### ***Clustering gerarchico***

Il *clustering* gerarchico (Jain et alii 1999) utilizza un metodo che consente di definire la similarità o dissimilarità di tutti gli elementi del set di dati. Il numero di *cluster* finali non è determinato a priori, inversamente in *clustering* di tipo non gerarchico come il K-Means e la sua versione fuzzy il numero di *cluster* deve essere imposto dall'utente prima che la procedura di *clustering* abbia inizio. Il metodo più popolare per scegliere a posteriori una suddivisione in classi è quello di generare un dendrogramma che permette di visualizzare a quale livello della procedura due elementi siano stati aggregati o suddivisi in base ad un criterio di similarità o dissimilarità.

Il *clustering* gerarchico può infatti essere sia di tipo agglomerativo che di tipo divisivo. Nel *clustering* agglomerativo ogni oggetto viene inizialmente posto nel proprio gruppo (che è ovviamente costituito dal solo oggetto stesso).

Successivamente i due gruppi più vicini vengono combinati in un unico gruppo. Nel *clustering* divisivo tutti gli oggetti vengono posti in un unico gruppo. I due oggetti che si trovano nello stesso gruppo ma risultano più lontani tra loro, vengono utilizzati come seme per i due nuovi gruppi. Tutti gli oggetti vengono posti quindi nel nuovo gruppo con il valore di distanza minore. La procedura continua iterativamente finchè non viene raggiunto un valore limite della distanza.

Per eseguire un *clustering* gerarchico è necessario determinare una misura della distanza tra due oggetti. Le distanze possono essere calcolate in vari modi, ed in uno spazio multivariato sono considerate misure di similarità o dissimilarità tra due oggetti. Ad esempio le distanze di tipo Euclideo determinano la distanza lungo una linea retta nello spazio n-dimensionale dove n è il numero delle variabili degli oggetti. In aggiunta un *cluster* di tipo agglomerativo necessita di un metodo per determinare la distanza tra i *cluster* che vengono man mano a formarsi:

Il *single linkage* utilizza la distanza minima tra ogni coppia di osservazioni prese all'interno di ogni *cluster*, quindi vengono uniti i *cluster* che presentano la distanza minima tra una qualsiasi delle coppie di osservazioni.

L'*average linkage* utilizza valori medi della distanza tra ogni coppia di osservazioni tra due *cluster*. Vengono quindi uniti i due *cluster* che presentano la minima distanza media calcolata.

Il *complete linkage* utilizza la massima distanza tra ogni coppia di osservazioni tra due *cluster* e vengono quindi uniti i *cluster* che presentano la minima distanza così calcolata.

Nel *Metodo di Wards* il criterio per l'unione di due *cluster* consiste nel minimizzare la somma totale delle deviazioni quadratiche medie dal centro del *cluster*.. Bisogna sottolineare che molte di queste tecniche necessitano del calcolo di una matrice di similarità/dissimilarità. Tali matrici occupano uno spazio pari a  $N(N/2-1)$  con N numero di vettori degli attributi. Una volta generato

un dendrogramma è possibile realizzare una divisione finale in *cluster* “tagliando” l'albero gerarchico ad un livello qualsiasi. A causa della grande quantità di memoria necessaria a memorizzare la matrice di distanza, i metodi gerarchici non sono utilizzati su set di dati di grandi dimensioni quali i dati *raster*. Basti pensare che la matrice di distanza di un semplice *raster* costituito da 100x100 celle e 3 livelli informativi occupa uno spazio di memoria pari a 45 milioni di elementi che se memorizzati con una precisione di 64 bit occupano oltre 3 gigabyte di memoria.

*Parte C*

*Elaborazioni sulle aree*

*test*

## **C.I Introduzione**

Sono state individuate tre differenti tipologie di analisi utili al raggiungimento degli obiettivi della ricerca di tesi. Con la cartografazione automatica si è cercato di affrontare il problema da due punti di vista differenti. Per l'individuazione di forme specifiche già a priori conosciute, si è sperimentato l'utilizzo del metodo della distanza statistica. Tale metodo è di tipo supervisionato, cioè prevede il riconoscimento e quindi la delineazione a priori di tali forme da parte dell'operatore. Naturalmente ci si era prefissi che tale fase non avrebbe dovuto comportare una cartografazione estensiva delle forme, ma la semplice delineazione di una piccola parte presente nell'area test. Questo metodo è stato utilizzato per individuare con successo le forme a controllo strutturale, sia attivo che passivo, quali versanti di faglia e versanti di linea di faglia sul raster A. Un secondo tipo di analisi sperimentata è quella della classificazione non supervisionata. Questa parte dell'elaborato ha consentito di verificare come una classificazione non supervisionata non sia facilmente implementabile come una di tipo supervisionato.

Le proprietà associate alle singole celle del raster non forniscono le informazioni necessarie a caratterizzare l'andamento della topografia di forme complesse la cui estensione spaziale richiede tecniche di analisi che tengano conto delle proprietà spaziali di tutte le celle che ricoprono la superficie della forma di interesse. Comunemente un'analisi di tale tipo si avvale di finestre di dimensioni fisse in termini di celle, la cosiddetta *neighborhood analysis*. La forma regolare della finestra imposta nelle tecniche di *neighborhood analysis* determina spesso risultati che non rispecchiano le reali caratteristiche delle forme del rilievo terrestre poiché queste sono caratterizzate da confini di forma irregolare. Ciò ha reso necessario sperimentare approcci alternativi.

Il presente lavoro propone un metodo per la classificazione delle forme del rilievo che si avvale del *clustering* iterativo applicato ai modelli digitali del

terreno, mirato ad evidenziare principalmente le forme descritte nella parte A.II.

Il metodo, attraverso opportune trasformazioni del set di dati iniziale, definisce nuovi attributi derivati dai parametri morfometrici iniziali. Tramite tale metodologia sviluppata ad hoc (Aucelli et alii 2005) ci si è anche posto l'obiettivo di realizzare in automatico una mappatura simile a quella ottenibile tramite metodi tradizionali, cercando quindi di simulare in maniera più efficace quelle fasi di generalizzazione che istintivamente la mente umana esegue quando il geomorfologo passa dalla lettura di una carta topografica al riconoscimento ed alla cartografazione delle forme di interesse. Il metodo è stata utilizzata per la classificazione del raster C ed del raster B.

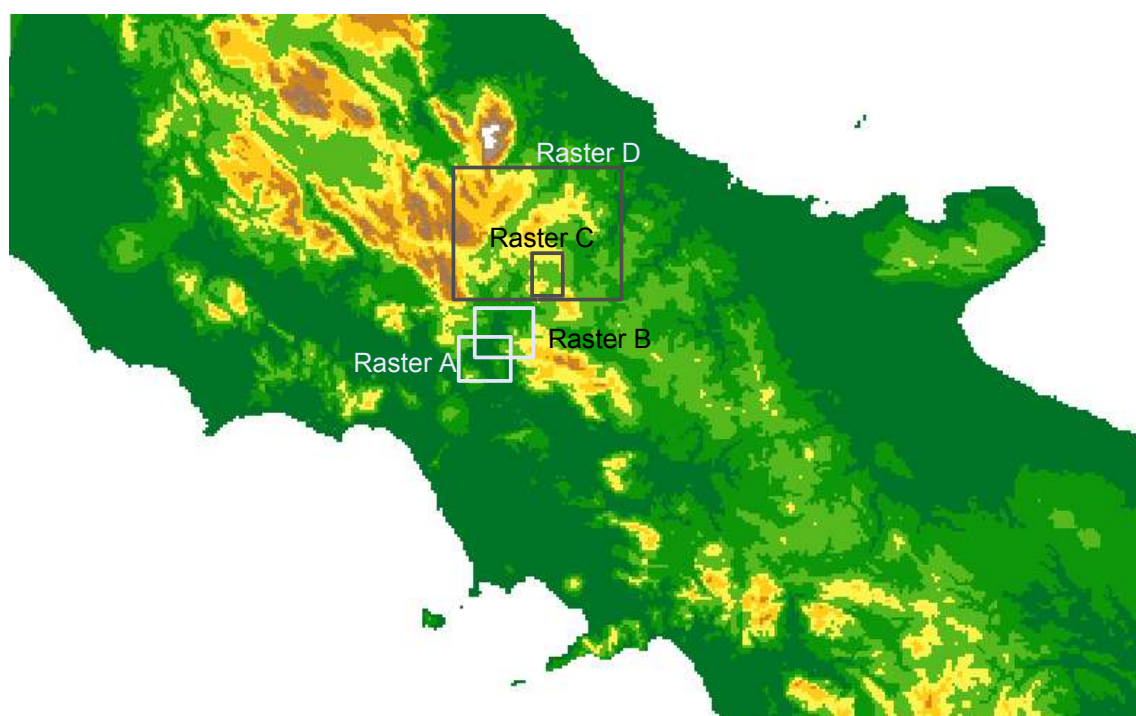
Sia per la classificazione supervisionata che per il *clustering*, di fondamentale importanza è stata la scelta della combinazione iniziale dei vettori degli attributi estraibili dal DTM tramite le tecniche descritte nel capitolo 3. Difatti, come descritto nel capitolo 5, è di fondamentale importanza la fase di selezione. Combinazioni differenti degli stessi portano a risultati completamente differenti utilizzando la medesima tecnica di classificazione. Le combinazioni scelte, sono frutto di una sperimentazione sì di tipo “tryal and error” ma comunque “guidata” da valutazioni a priori. Difatti, sia la metodologia di classificazione scelta che la combinazione “tecnica di classificazione-vettori degli attributi” è stata guidata dall'analisi delle basi teoriche descritte nella parte B.

Per la correlazione di superfici relitte è stata sperimentata una metodologia diversa, basata su di un algoritmo scritto appositamente. Questo obiettivo risulta essere differente da quello della individuazione delle varie tipologie di forme. Infatti il criterio geomorfologico non è più l'interpretazione della forma del versante ma l'interpretazione della successione di generazioni delle superfici.

Il criterio è puramente geometrico visto l'impossibilità di confermare con dati di altra natura l'effettiva età delle superfici. L'ordinamento è stato realizzato per le superfici terrazzate del raster A e del raster D, che comprende le superfici di glacis.

I DTM in formato raster sono stati realizzati tramite digitalizzazione delle linee di livello di tavolette topografiche in scala 1:25.000. Oltre alle linee di livello sono stati digitalizzati anche diversi punti quotati in maniera da evitare il cosiddetto effetto di “terracing” (cfr. paragrafo 3.7). Le linee di livello ed i punti quotati sono stati digitalizzati tramite ArcView 3.2 ed importati nel software GIS open source GRASS (GRASS 2002) versione 5.0.1 sotto forma vettoriale.

Qui, tramite l'algoritmo s.surf.rst, è stata realizzata l'interpolazione generando un DTM con risoluzione di 40 metri per tutte le aree. Il suddetto algoritmo è stato impiegato più volte per le stesse aree sperimentando diversi parametri di interpolazione, a causa della impossibilità di conoscere a priori i parametri ottimali in grado di limitare l'effetto di “terracing” e “ringing” oltre che la eventuale comparsa di scalettature dovute alla segmentazione (Cebecauer et alii, 2002). Utilizzando i parametri di tensione=18.0, smoothing=0.1, segmax=400 ed npmin=800 si è ottenuto un DTM considerato qualitativamente accettabile per le analisi dei parametri morfometrici.



*Fig 6.1 Localizzazione delle aree test*



## **C.II *Le aree test***

### **6.2 L'alta valle del fiume Volturno.**

Il territorio in esame è situato tra gli abitati di Isernia e Roccaravindola. L'area è stata suddivisa in due raster al fine di sperimentare differenti tecniche di analisi mirate all'individuazione di specifiche morfologie. Il raster che dell'area meridionale comprende la quasi totalità della carta geomorfologica dell'alto bacino del Volturno, redatta da Brancaccio et alii (2000). Il raster A è stato quindi utilizzato principalmente per sperimentare una tecnica volta all'individuazione degli ordini di terrazzi, anche se in fase di studio si è deciso di sfruttare l'area anche per una classificazione supervisionata per individuare i versanti di faglia. Il raster B corrisponde ad un segmento del bacino posto più a nord ed è stata utilizzata per classificare l'area tramite la tecnica di *clustering* descritta nel capitolo 7.

L'alta valle del Volturno è caratterizzata da terreni calcarei e silicoclastici di età meso-cenozoica e subordinariamente da terreni quaternari di origine alluvionale, rilevabili lungo i principali assi drenanti. In particolare i depositi quaternari presenti nell'area analizzata risultano inseriti in una ampia depressione intermontana di origine tettonica, riconducibile alla dinamica trascorrente attiva tra il Pliocene ed il Pleistocene inferiore (Di Bucci et alii 1999) e sono caratterizzati da facies di tipo fluvio-lacustre. La piana risultante (sistema della Piana di Isernia-Venafro) risulta bordata ad est e sud-est dal complesso carbonatico della Montagnola di Frosolone e a nord dal complesso Flyscoide dell'Unità di Agnone. La Piana di Isernia si collega a sud con la piana di Venafro. Le formazioni terrigene meso-cenozoiche affioranti nell'area sono comunemente identificate come Unità Molisane (Patacca et alii 1992b) la cui origine viene attribuita ad un dominio esterno a quello della piattaforma del Matese. I depositi silicoclastici dell'Unità di Agnone sono estesi nell'area a nord-est della piana di Isernia e sono costituiti dai tre membri identificati dal

basso verso l'alto come membro Poggio-Villanelle, membro Verrino e membro Sente. La successione comprende rispettivamente strati arenitici fini che passano a strati con elementi più grossolani nel membro intermedio caratterizzato oltre che dal maggior spessore dei depositi arenitici, da intercalazioni di biocalcareniti e calciruditi risedimentante, da blocchi calcarei sparsi e da olistoliti di grandi dimensioni di calcari a briozoi e litotamni, calcari di piattaforma a rudiste, calcareniti a frammenti di rudiste, calcareniti a lepidocicline e calcari marnosi ad *Orbulina*. L'ultimo membro, il membro Sente, è costituito principalmente da argille e marne grigie con intercalazioni centimetriche di arenarie fini. Nell'area in esame, la parte della successione affiorante dell'Unità di Frosolone è rappresentata dalla formazione di Longano, costituita da calcitorbiditi del Tortoniano superiore.

Come già accennato, all'interno del sistema della piana di Isernia-Venafro sono stati individuati depositi riconducibili a differenti fasi del riempimento quaternario. Questi si distinguono in un'Unità principale ed in un'Unità di fondovalle (Brancaccio et alii 2000) ulteriormente suddivisibili in membri con differenti caratteristiche lito-stratigrafiche ma che spesso presentano anche caratteristiche morfologiche distinte. L'Unità principale è la più estesa e costituisce il primo evento deposizionale riscontrabile nell'area. E' costituito da depositi ghiaiosi, limosi ed argillosi di facies fluviale. All'interno di tale unità si riscontrano anche depositi travertinosi, che costituiscono ad esempio la placca sommitale dell'antico terrazzo fluviale su cui sorge l'abitato di Isernia dove sorge l'abitato di Isernia, dove costituisce una placca sommitale al di sopra di un antico terrazzo fluviale. L'Unità di fondovalle è costituita prevalentemente da ghiaie poligeniche e si incastra nell'Unità principale. I terrazzi associati alle unità distinte sono sia di tipo erosionale che deposizionale. Sono individuabili almeno quattro ordini di terrazzi variamente distribuiti. Il primo e più antico ordine (I ordine) è di tipo deposizionale ed è impostato sull'Unità principale. Questo ordine risulta essere il più esteso ed è diffusamente disseccato dai tributari del Volturno rendendo molto articolata

la distribuzione planimetrica dei suoi lembi residui. La sua estensione altimetrica è molto ampia e varia da 490 m s.l.m. a circa 250 m s.l.m. seguendo approssimativamente un gradiente NO-SE. Nell'area di studio sono presenti sia conoidi di deiezione a contatto con i rilievi carbonatici sia vere e proprie conoidi alluvionali, alcune di notevole estensione, i cui depositi contribuiscono ad aumentare l'acclività della morfologia dei lembi di terrazzo (foto 2). Si riconosce una fase di tettonica estensionale a partire dal Pleistocene medio che ha influenzato l'evoluzione della piana alluvionale e quindi delle forme terrazzate. Risulta evidente, in particolare, la duplicazione del primo ordine evidenziata dalla presenza di lembi dislocati in prossimità dell'abitato di Isernia ove è possibile riconoscere vari allineamenti di faglia. Nonostante siano oramai scomparse le evidenze dirette di attività recenti lungo i principali lineamenti tettonici noti in letteratura (vedi di Bucci et alii 1999) - se non in un caso nei pressi dell'abitato medesimo ove è possibile osservare un versante di faglia - di attività recenti lungo i principali lineamenti tettonici noti in letteratura (Bucci et alii 1999) sono presenti numerose evidenze indirette di come tale attività sia continuata ed abbia condizionato l'evoluzione della piana alluvionale. La presenza di faglie è testimoniato dallo sviluppo del reticolo idrografico che assume un pattern angolato, dalla presenza di gomiti ad "uncino" lungo il Volturno ed il Cavaliere che non seguono la linea di massima pendenza, e dalla posizione anomala dell'alveo del F. Cavaliere rispetto al suo fondovalle che ne fa presumere una migrazione recente per basculamento della valle. E' quindi ipotizzabile che parte dei lembi riscontrati, anche laddove non siano evidenti scarpate di faglia, risultino dislocati, anche se è oramai scomparsa l'evidenza morfologica. Caso emblematico è il terrazzo su cui poggia l'abitato di Isernia, la cui pendenza varia considerevolmente a partire dalla confluenza dei due torrenti che lo bordano (Carpino e Cavaliere) fino al raccordo, verso monte, con la piana sovrastante (Località Le Piane) che non è giustificabile con fenomeni di *slope decline* del terrazzo. Sono state difatti segnalate durante lavori di sbancamento per la costruzione della superstrada Napoli-Vasto, lungo il banco di travertino al

top del substrato argilloso, faglie dirette con rigetto decimetrico.

I terrazzi di II ordine sono prevalentemente di origine erosionale, intagliati nei terrazzi di I ordine. In particolare risultano molto estesi nella parte meridionale della piana di Isernia, ove non sono più riscontrabili terrazzi di I ordine veri e propri ma risultano invece presenti dei corpi di conoide erosi al piede ove si incastra il secondo ordine. Anche i terrazzi di II ordine risultano essere rappresentati da conoide tagliate a loro volta al piede dalle successive fasi erosionali ed a contatto deposizionale con i terrazzi di III ordine. La loro estensione altimetrica è compresa tra i 380 m.s.l.m ai 210 m.s.l.m. L'Unità principale, e con essa il I e II ordine di terrazzi, scompare in prossimità della piana di Venafro, all'altezza della conoide di Località Carpinete. Tale fatto è presumibilmente da attribuire ad un dislocamento tettonico lungo la linea di faglia "Pozzilli-Capriati" che ha determinato il sollevamento relativo della piana di Isernia, e di conseguenza un aumento dei processi di incisione ed erosione fluviale in tale piana mentre nell'area ribassata ha prevalso la deposizione di facies fluvio-lacustri. Le due piane risultano attualmente completamente aggradate e l'assenza del I e II ordine nella piana di Venafro è riconducibile al fatto che il dislocamento ha avuto luogo dopo la messa in posto e la prima fase di incisione dell'Unità Principale. I terrazzi di III e IV ordine sono deposizionali e costituiscono l'Unità di fondovalle. Risultano scarsi i lembi di III ordine mentre i lembi di IV ordine sono molto estesi e costituiscono l'attuale piana di esondazione del torrente Vandra, Cavaliere e del fiume Volturno. Lungo la piana di Isernia le quote dei due livelli di terrazzo variano da un massimo di 290 m.s.l.m ad un minimo di 200 m.s.l.m. Le forme di origine strutturale sono ben presenti nell'area di studio a causa dei numerosi lineamenti che la attraversano. Si individuano versanti di faglia ampiamente evoluti, retrocessi per *slope replacement*, lungo i margini degli affioramenti terziari, talora degradati a faccette trapezoidali o triangolari (foto 1). Nell'area sono altresì individuabili alcune superfici relitte, di formazione anteriore alla dislocazione che ha generato

la depressione tettonica in cui è iscritto il bacino, la cui estensione varia da piccoli lembi, presumibilmente singoli ordini variamente smembrati e dislocati altimetricamente, fino a superfici abbastanza estese lungo i margini dei principali versanti di faglia. L'attribuzione di queste superfici a particolari ordini di generazioni risulta difatti particolarmente difficoltosa per cui non sono state considerate come elementi ideali per testare l'algoritmo sviluppato e descritto nel paragrafo 7.4.



*Foto 1 - Versante di faglia dissecato e rimodellato a faccette trapezoidali (località Carpinete)*

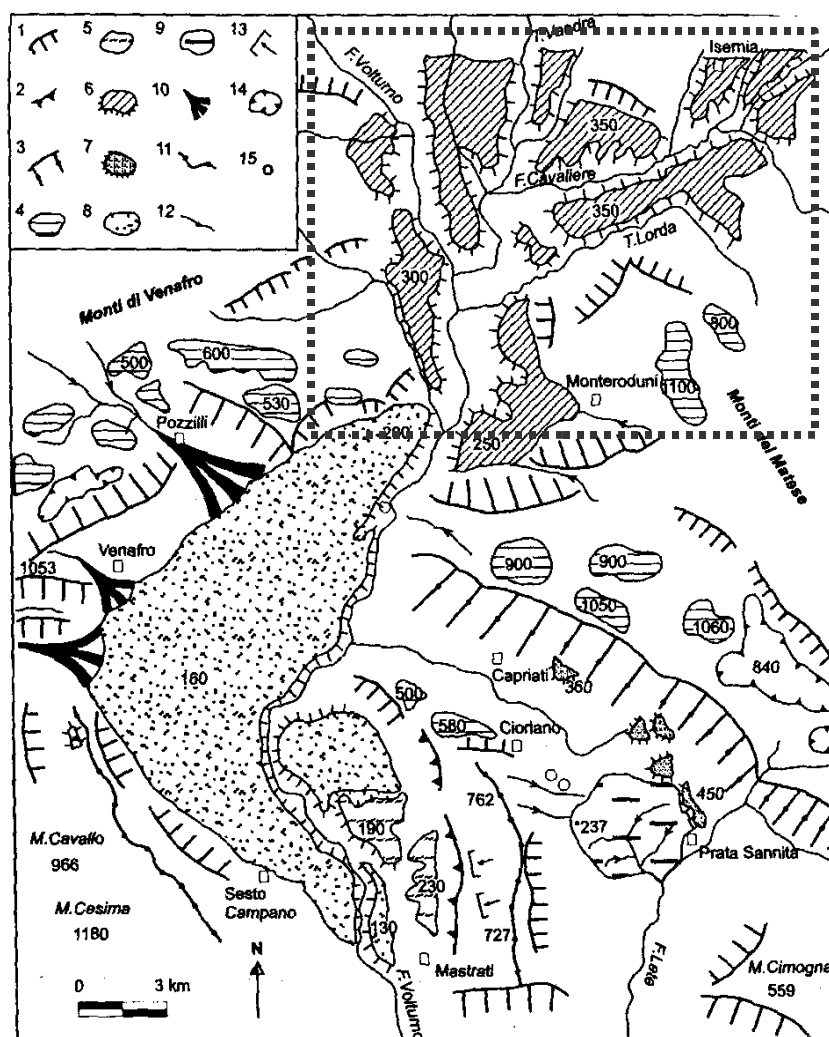
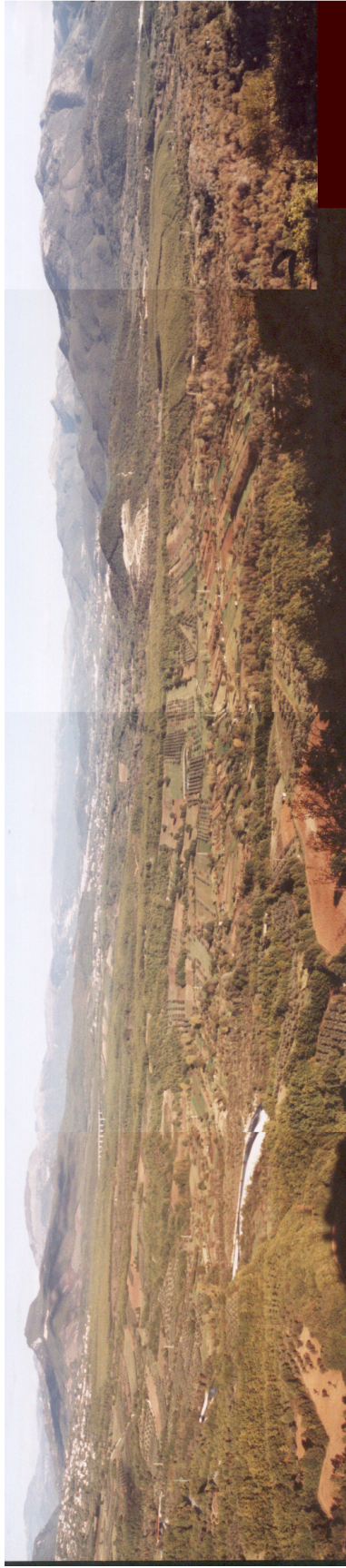


Fig 6.2 Schema geomorfologico dell'area dei terrazzi del fiume Volturno 1) versanti di faglia 2) versanti di linea di faglia 3) versanti con indizi di riattivazione tettonica quaternaria 4) lembi di paleosuperfici 5) terrazzi di erosione su flysch 6) terrazzi deposizionali ed erosionali di vario ordine (bacino di Isernia) 7) terrazzi deposizionali su materiale di conoide 8) fondovalle della piana di Venafrò 9) depositi limo-vulcanici terrazzati 10) conoidi 11) linea di cresta a controllo strutturale 12) corsi d'acqua susseguenti 13) valle sospesa 14) polje 15) sorgenti sulfuree. (Da Brancaccio et alii 2000), il riquadro tratteggiato indica l'area ricoperta dal raster A.



A)



B)

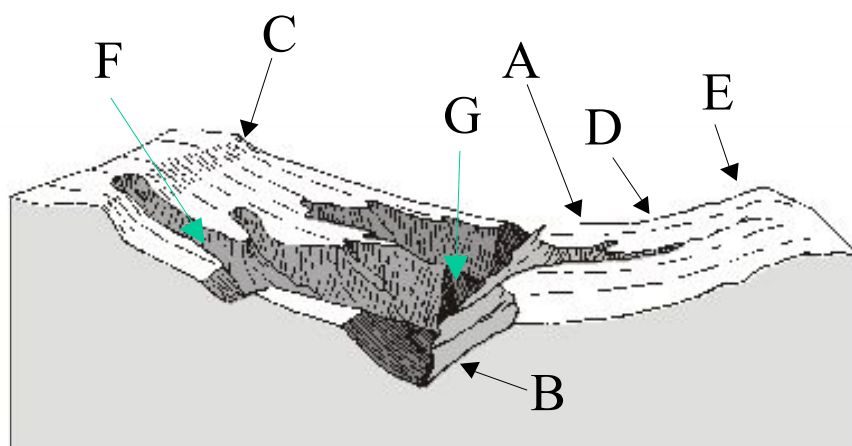
Foto 2 Complesso dei terrazzi fluviali della valle del fiume Volturno. A) Area a nord della confluenza con il fiume Cavaliere, caratterizzata da ampi lembi di I ordine. B) Area a sud della confluenza con il fiume Cavaliere: i depositi associati ai terrazzi di I e II ordine si rinvengono sotto forma di conoidi alluvionali

### 6.3 La valle del fiume Trigno

Il fiume Trigno è caratterizzato da una lunghezza di 120 km ed il suo bacino di cattura è di circa 1200 kmq con una quota media di circa 600 m.slm. Le formazioni geologiche che affiorano lungo la media valle sono di età Oligocenica-basso Messiniano (Ortolani et alii 1975, Lentini 1979, Patacca et alii 1992b) e possono essere ascritte alle sequenze molisane in parte già descritte per la valle del Volturno (unità di Frosolone-Agnone e Tufillo-Daunia). Le unità Sannitiche e la formazione di S. Bartolomeo. Dal punto di vista litologico tali formazioni hanno carattere prevalentemente pelitico (argille varicolori, marne ed argille, argille e sabbie del Plio-Pleistocene ecc..) e che determinano una bassa resistenza all'erosione dei versanti. Sono presenti anche formazioni caratterizzate da litologie più resistenti, per la maggior parte intercalate alle prime, quali calcari e marne ed arenarie. Le relazioni tettoniche e stratigrafiche nell'area di studio sono piuttosto complesse caratterizzate da numerosi sovrascorrimenti e faglie di varia tipologia. La maggior parte dei lineamenti tettonici sono orientati in direzione W-E, NW-SE e NE-SW. Risulta chiaro dall'andamento del reticolo come il suo sviluppo si sia stato influenzato dalla presenza di tali lineamenti. L'area del bacino idrografico del fiume Trigno è coperta dal raster D. Il Trigno è uno dei tre principali corsi d'acqua che drenano la parte est dell'Appennino Molisano in direzione ovest-est e che sfociano nel mar Adriatico tagliando attraverso *trusth* e strutture a falde caratteristici di questa porzione della catena. Nella sua percorso iniziale, il reticolo è impostato su di un sistema vallivo dominato da forme mature. Nella zona centrale le forme vallive sono caratterizzate da un'ampia fascia che ha subito un generale sollevamento durante il periodo Plio-Quaternario. Qui le forme assumono caratteristiche che evidenziano una più recente evoluzione del reticolo che risulta confinato in strette valli a V con fianchi vallivi caratterizzati da elevata pendenza e forme che indicano una generale instabilità dei versanti. Queste caratteristiche sono riscontrabili anche nella la parte più alta della valle del fiume Volturno pre-



cedentemente descritta (torrente Vandra e Vandrella) che risulta caratterizzata



*Fig 6.3 Forme caratteristiche dei fianchi vallivi del F. Trigno. A) Superficie a basso angolo di pendenza interpretabile come glacis d'erosione in roccia tenera; C) Cornice sommitale interpretabile come risultato dell'arretramento del versante durante la formazione del glacis; D) Forma concava ed E) Forma convessa che raccordano il glacis A allo spariacque; B) Forma rettilinea (scarpata fluviale) che raccorda il glacis al fondovalle; G) Faccetta di interfiumo; F) Frana. (Da Aucelli, 1999 modificato)*

dalla la presenza di numerosi fenomeni franosi ed alvei confinati in strette valli a V. Buona parte delle caratteristiche litologie e morfologiche risultano tra queste due aree risultano infatti molto simili. Lungo i bacini dei tre fiumi a vergenza adriatica, contrariamente alla valle del Volturno, i lembi di terrazzi fluviali sono scarsamente presenti e si rinveno raramente lungo il fondovalle del corso d'acqua principale, come lembi isolati e di modesta estensione, è più estesamente nella parte finale in prossimità della foce dove è possibile osservare due o tre ordini con dislivelli fino a dieci metri rispetto al fondovalle. In particolare, nella valle del Trigno, è possibile individuare solo lembi nastriformi di scarsa estensione. E' quindi ipotizzabile che durante l'evoluzione del sistema vallivo questo non abbia subito rare ed irregolari periodi caratterizzati dalla prevalenza all'accumulo, e che l'attuale assetto geomorfologico è dovuto ad una forte prevalenza al trasporto lungo la parte mediana della valle degli apporti laterali con brevi fasi. Le modeste forme di accumulo sono state rimodellate o cancellate dalla tendenza generale all'approfondimento.

Nelle valli a V, i fianchi vallivi sia del Trigno che dei suoi tributari, sono caratterizzati da profili che mostrano una sequenza composita (fig 6.3). La parte più alta è caratterizzata da segmenti dalla forma concavo-convesse poco accentuata che sormontano superfici più rettilinee che presentano pendenze variabili tra i 3 e gli 8 gradi, interpretabili come relitti di superfici di glacis. Queste sono separate fondovalle da scarpate di origine fluviale ad elevata pendenza e con dislivelli fino a 20 metri. Le superfici a pendenza più blanda possono essere interpretate come superfici erosionali formatesi durante periodi caratterizzati da bassa energia del reticolo dovuti a minore apporto idrico o a maggiore apporto detritico. Sono prevalentemente riscontrabili in rocce tenere quali marne sottilmente stratificate, argilliti e argille caotiche e vengono ascritte a quel tipo di superfici note in letteratura come “glacis en roche tendres” (Birot et alii 1958, Vaumaz 1963). Verranno qui di seguito definite più brevemente come “Glacis Valli Sospesi” o GVS (Aucelli 1999, Aucelli et alii 2001) poiché si rinvenivano prevalentemente a diversi metri di altezza rispetto al fondovalle attuale. Spesso tali superfici risultano fortemente dissecate tanto da risultare in lembi isolati. La dissezione porta alla perdita di raccordo con il fondovalle i cui fianchi risultano ridotti a faccette triangolari sormontate da creste che fanno da spartiacque tra le vallecole che separano i vari lembi di glacis.



*Foto 3 - Glacis Vallivo Sospeso alla confluenza del torrente Vella  
(da Aucelli 1999)*

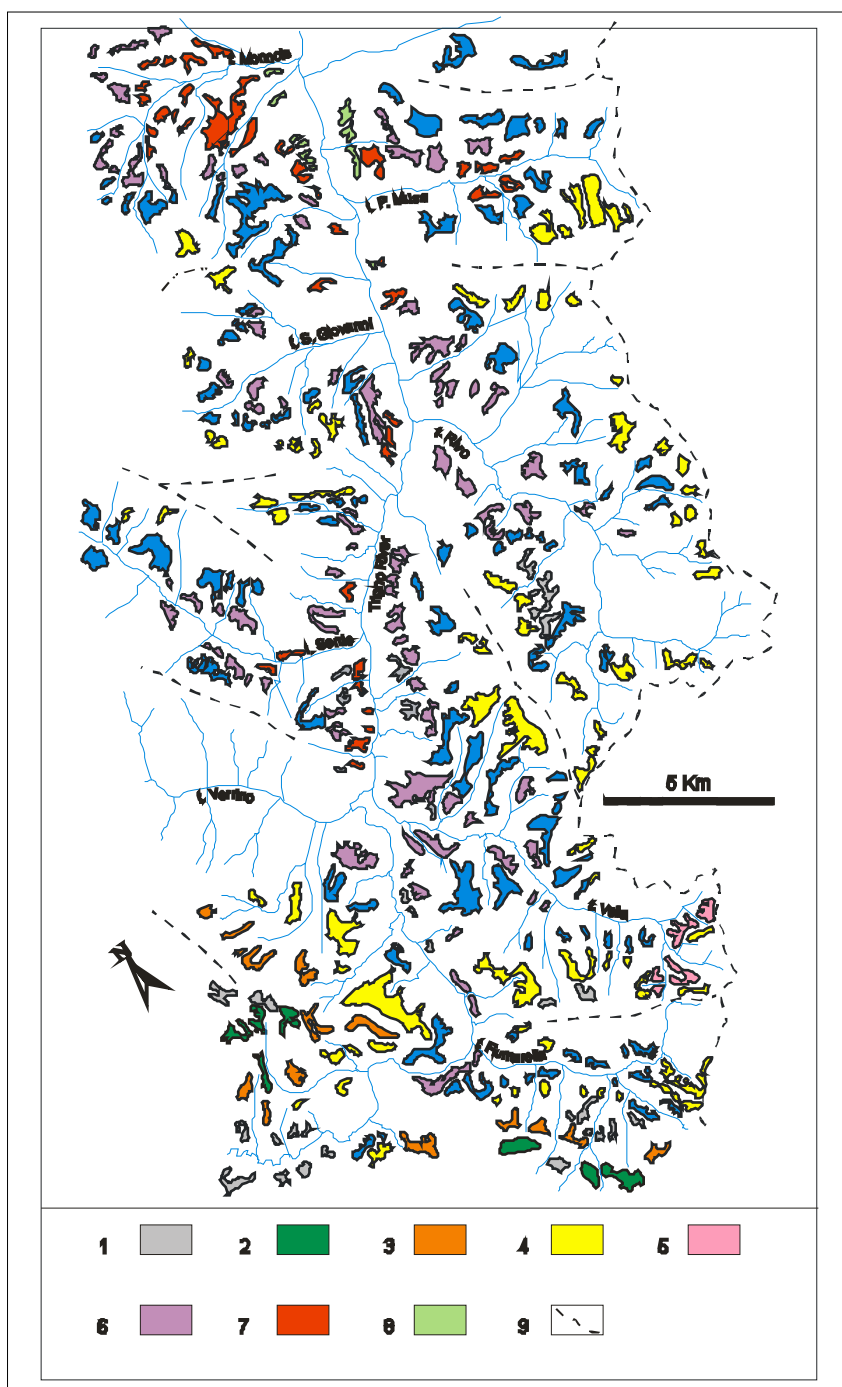


Fig 6.4 Mappa della distribuzione dei glacis all'interno della media valle del Trigno. 1) glacis di incerta sede; 2-8) glacis suddivisi per ordine, dal più antico al più recente; 9) spartiacque principale e secondari. (Da Aucelli 1999) Il raster C ricopre l'area del bacino del T.Vella mentre il raster D l'intera mappa rappresentata.

### **C.III *Elaborazioni sui raster***

#### **7.1 Precedenti conoscenze**

La maggioranza delle forme naturali sono geometricamente complesse specie se derivano dall' azione di più processi. Inoltre la maggior parte degli elementi del paesaggio non sono caratterizzati da confini ben definiti presentando spesso delle zone marginali di transizione con altri elementi. Elementi del paesaggio possono essere visti anche da un punto di vista gerarchico, con forme di varia genesi, riconoscibili ad una determinata scala di analisi, incluse in più ampi elementi cartografabili a scala più piccola. Ad esempio un versante di piano di faglia, forma caratteristica dei nostri paesaggi appenninici, fa parte – nel senso di essere compreso nei confini cartografabili - di un più ampio elemento del paesaggio che può essere un versante vallivo o lo stesso orogene. Una valle può essere suddivisa nei suoi vari elementi o essere cartografata come elemento singolo dipendentemente da come il geomorfologo intende delineare gli elementi del paesaggio. Non si tratta difatti solo di un problema di scala ma anche di scelta degli elementi di interesse. E' chiaro che anche ad una scala di analisi che consente di identificare elementi non più riconoscibili a livelli di dettaglio inferiori, il geomorfologo tende a definire i confini delle forme con maggior accuratezza ma sempre utilizzando dei criteri personali e comunque finalizzati all'utilizzo per la quale la cartografazione viene realizzata.

Elementi non di interesse, anche se ben evidenziabili, potrebbero essere non cartografati o inglobati in generici confini (si pensi come esempio alla delineazione dei versanti in frana invece che i singoli corpi franosi lungo il versante). Anche nel caso sia chiaramente definita a priori la finalità ed il tipo di forme da cartografare è possibile che più geomorfologi siano chiamati a contribuire alla realizzazione dell'elaborato. In tal caso risulta maggiormente evidente il problema della soggettività. I confini delle forme saranno mappati in maniera

differente a seconda della propria esperienza e interpretazione. Ciò non solo perché la capacità di riconoscere forme di genesi diversa dipende da quanti modelli interpretativi possiede il geomorfologo, ma anche perché la stessa fase di analisi della superficie topografica (sia essa fatta sul campo, che su carte o su foto aeree, etc.) viene svolta con intensità ed accuratezza variabili da persona a persona. Per non dire che addirittura lo stesso geomorfologo, quando lavora a lungo su aree vaste, può capitare di usare criteri di analisi e livelli di attenzione un po' diversi da luogo a luogo. Si sono vagliate quindi diverse ipotesi di lavoro al fine della sperimentazione ed implementazione del metodo.

### ***Metodi combinatori***

Nel mondo anglosassone il metodo di Hammond (1964a) ha suscitato molti consensi. Si tratta di un metodo di classificazione quantitativo di tipo gerarchico. L'autora utilizza una finestra di analisi quadrata di 9,65 km di lato, per regionalizzare alcuni parametri morfometrici definiti a priori. I parametri utilizzati sono la percentuale di superficie a bassa pendenza, 8% *gentling sloping* secondo l'autore, il *local relief*, e la distribuzione delle pendenze sopracitate lungo il profilo, utilizzando rispettivamente 4, 6 e 4 classi. L'analisi viene condotta per blocchi non sovrapposti contraddistinti dunque da una combinazione unica di tali parametri. I criteri utilizzati danno un totale di 96 combinazioni teoriche possibili, che l'autore ordina gerarchicamente in classi e sottoclassi di cui 5 classi dominanti: pianure, altopiani, colline e montagne, colline e montagne "aperte", cioè caratterizzate da porzioni di superficie che si raccordano alle pianure. Questo metodo manuale è stato successivamente automatizzato da Dikau et alii (1995).

In particolare Dikau impiega per estrarre i parametri di Hammond, una finestra mobile, impiegando la tecnica della *neighborhood analysis*. I valori estratti sono estratti da un DTM con 200 metri di risoluzione e vengono calcolati tramite una finestra di analisi di 49x49 celle e quindi associati alla cella focale.

Le classi sono ordinate utilizzando il criterio di classificazione originale di Hammond. Tale metodo è stato utilizzato per realizzare una classificazione del DTM del New Mexico e paragonato all'originale classificazione manuale di Hammond della stessa area, notando, come prevedibile dal differente schema di estrazione dei parametri, differenze nel numero di sottoclassi ottenute e loro distribuzione spaziale. Questo tipo di classificazione automatica è facilmente riproducibile ma si adatta bene ad una classificazione del territorio a scala molto piccola, evidenziando le caratteristiche del macrorilievo (montagne, altipiani ecc.). Inoltre risulta un metodo molto formalizzato in quanto le classi sono predefinite ed evidentemente adattate alle specifiche caratteristiche dell'area di studio analizzata (New Mexico).

Tecniche combinatorie sono state utilizzate ampiamente nel campo dell'analisi della pericolosità in aree soggette a fenomeni franosi (Carrara et alii 1995). Si tratta di tecniche mirate all'identificazione delle aree a maggiore suscettibilità tramite analisi statistiche basate sul criterio di Bayes. In pratica dati due o più livelli informativi (pendenza, litologia ecc.) cartografati e riclassificati (ad esempio la pendenza può essere riclassificata in pochi intervalli rispondenti alle maggiori classi di pericolosità note in letteratura, le formazioni geologiche accorpate in unità litologiche ecc..) si effettua una combinazione tramite *map overlay* risultante in una carta formata da vari poligoni di intersezione detti “Unique Condition Area” o UCU (Carrara et alii 1995). Ogni poligono è caratterizzato da una combinazione di parametri “unica”. I poligoni vengono quindi analizzati con tecniche dipendenti dalla finalità dello studio, utilizzando queste combinazioni di parametri. Tra i primi tentativi di utilizzare metodi combinatori citiamo Speight (1967) che utilizza uno schema combinatorio manuale differente rispetto all'approccio tassonomico di Hammond. Speight non effettua una classificazione tassonomica delle combinazioni risultanti e non utilizza finestre regolari per regionalizzare i parametri morfometrici. L'approccio è invece del tutto simile a quello utilizzato nei lavori che utilizzano la tecnica

delle UCU. Difatti egli originariamente utilizza quattro livelli informativi – quota, pendenza, curvatura planare, curvatura tangenziale ed unità di bacino, riclassificati in pochi intervalli in modo da poterli combinare tramite intersezione manuale. Successivamente i poligoni di intersezione vengono raggruppati in poche classi in base a criteri di tipo soggettivo. Il metodo proposto da Speight si distingue anche per la scala di analisi utilizzata. Gli accorpamenti sono interpretati infatti come elementi più caratteristici del mesorilievo, quali creste, interfluvi, fianchi vallivi ecc... Il metodo combinatorio, indipendentemente da come poi si procede all'individuazione delle classi, porta ad un'elevato numero di combinazioni uniche. Statisticamente molte delle possibili combinazioni teoriche non si presentano o risultano in poligoni di estensione talmente ridotta da non essere di fatto statisticamente significanti. Ad esempio nel lavoro originale di Speight molte delle 432 possibili combinazioni uniche, non si realizzano. Analogamente si può riscontrare nei lavori di Hammond e Dikau che alcune combinazioni risultano presenti ma ricoprono una percentuale ridottissima dell'area analizzata, o non sono affatto presenti. Nonostante diverse discrepanza tra le “unità di paesaggio” o “elementi del pendio” individuati da Speight in questo modo ed il raffronto con carte tematiche realizzate tramite l'analisi di foto aeree e carte topografiche, Young (1972) definisce tale approccio come “di notevole interesse e meritevole di ulteriori sviluppi”. Difatto l'accorpamento delle UCU effettuato manualmente implica un notevole grado di soggettività.

### ***Metodi basati sulla cluster analysis***

I metodi basati sulla *cluster analysis* impiegano algoritmi di tipo iterativo quali il k-means, ISODATA, o il fuzzy c-means. La caratteristica principale di tali metodi di classificazione riguarda la possibilità di includere numerosi parametri morfometrici estratti a livello della singola cella del DTM raster. Spesso la scelta dei parametri varia da autore ad autore. Si utilizzano strati informativi derivati sia dall'analisi differenziale che dalla *neighbourhood analysis* e dall'analisi del flusso

idrico. I più comuni impiegati, anche nei lavori precedentemente citati nel capitolo 5, sono solitamente il *Local Relief*, il *wetness index*, la pendenza e le curvature planare (oppure quella tangenziale) e del profilo (cfr. capitoli 3 e 4). Le metodologie, come descritto in 5.4, portano ad una suddivisione in classi che necessita di un'interpretazione finale delle stesse ma l'intero processo di suddivisione è completamente quantitativo, benché non manchino in letteratura approcci semi-quantitativi che combinano la classificazione non supervisionata all'analisi visiva di mappe derivate (Guzzetti & Reichenbach 1994).

## **7.2 Il metodo sviluppato per la classificazione non supervisionata**

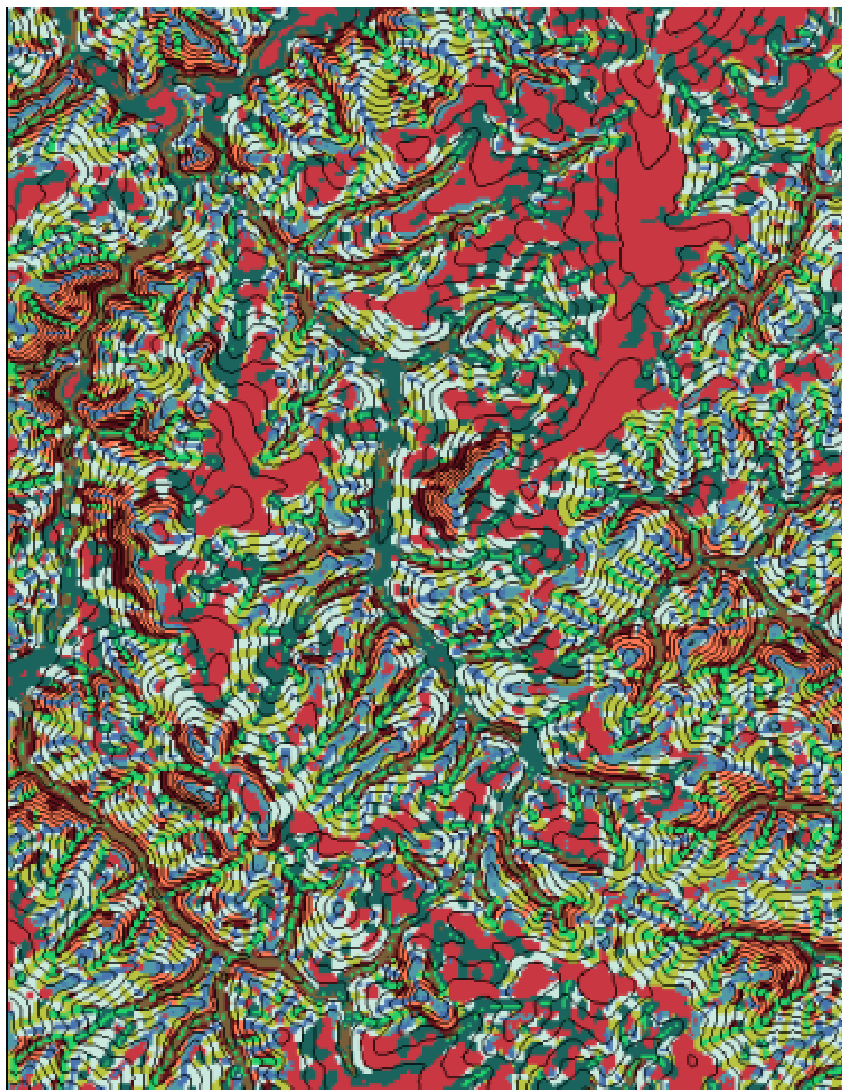
Un metodo di classificazione non supervisionata è stato sviluppato allo scopo di superare i limiti riscontrati inizialmente nell'utilizzo delle classiche procedure di *clustering* iterativo applicate alla classificare dei DTM tramite parametri morfometrici e topografici. Una delle principali motivazioni che hanno portato all'implementazione di un nuovo metodo è stata la assenza di una metodologia contemporaneamente valida e computazionalmente poco costosa e che inoltre tenesse conto della sovrapposizione delle variabili morfometriche nello spazio geografico. Ad esempio algoritmi di classificazione quali il *fuzzy c-means* tengono conto della sovrapposizione degli attributi nello spazio multivariato (cioè lo spazio degli attributi) ma producono dei risultati che non rispecchiano i confini naturali delle forme del rilievo nello spazio geografico. Tale approccio verrà di seguito definito “pixel based”. Il termine si riferisce al fatto che gli attributi utilizzati dall'algoritmo di classificazione, qualsiasi esso sia, derivano da proprietà associate al singolo pixel. Alcune tecniche di estrazione dei parametri morfometrici in effetti utilizzano le proprietà dell'intorno della cella. Questi valori sono però caratteristiche della singola cella. Un tipo di classificazione non supervisionata “pixel based” è visibile in figura 7.1. Dal raster C sono stati estratti tre parametri morfometrici: curvatura del profilo, curvatura tangenziale e



pendenza. Tali proprietà sono caratteristiche locali della superficie. I tre raster sono stati impiegati per suddividere l'area in 10 classi tramite l'algoritmo k-means. Un risultato del genere si può ottenere facilmente utilizzando la seguente sequenza di comandi della shell di GRASS che impiega il modulo *r.slope.aspect* e ed i due moduli, *r.clustering* ed *r.fuzzy*, scritti ed implementati durante lo svolgimento dell'elaborato di tesi:

```
>r.slope.aspect input=dtm_area slope=pendenza profc=curvatura.p  
tangc=curvatura.t  
>r.clustering input=pendenza,curvatura.p,curvatura.t output=raster_risultato k=10  
maxiter=100
```

Il risultato tuttavia non può essere considerato soddisfacente poiché i tre layer utilizzati non vengono elaborati tenendo in considerazione l'adiacenza spaziale dei pixel e quindi l'insieme dei pixel di forme composite può venire facilmente suddiviso in differenti *cluster*.



*Fig 7.1- Classificazione dell'area ricoperta dal raster C utilizzando un metodo "pixel based". Le classi ottenute risultano di difficile interpretabilità*

Un certo grado di aderenza a tale tipo di necessità è data dall'inclusione, come visto nel paragrafo precedente, di parametri morfometrici estratti a varie scale di analisi tramite metodi ai minimi quadrati quali quelli proposti da Evans (1972) o Zevenbergen e Thorne (1986) ed ulteriormente indagati ed implementati da Wood (1996) in un modulo del Gis GRASS chiamato `r.param.scale`. E' bene spiegare che l'estrazione dei parametri in questo caso avviene tramite una finestra di analisi all'interno della quale viene interpolata una superficie quadratica tramite il metodo dei minimi quadrati (cfr. capitolo 3). L'interpolazione si avvale dei valori delle celle all'interno della finestra, calcolando il valore per la sola cella

centrale (cella focale) del parametro desiderato (quota, pendenza, curvature ecc..). Essendo la finestra di forma regolare, non rispecchia l'andamento naturale dei bordi degli elementi del paesaggio. La superficie quadratica interpolata, data la sua semplicità, risulta difatti influenzata dai valori di elementi adiacenti ricadenti all'interno della finestra di analisi. Maggiore è la scala di analisi (ovvero la dimensione della finestra utilizzata) più celle estranee all'elemento corrispondente ad una particolare forma verranno inglobati nel calcolo della superficie interpolante e conseguentemente dei parametri derivati. Un metodo per poter cercare di superare tale difetto potrebbe essere, in linea teorica, quello di utilizzare un gran numero di finestre di dimensioni variabili ( $3 \times 3, 5 \times 5, 7 \times 7 \dots N \times N$ ), come sperimentato da Brown et alii (1998), in maniera tale da estrarre dal DTM un set di variabili, che rappresentano lo stesso attributo ma estratte a scala di analisi differente, che forniscano quindi informazioni su come varia la forma dell'intorno di ogni cella al variare della scala. Tuttavia all'aumentare della dimensione della finestra si evidenzia velocemente il trend della superficie realizzando di fatti un'analisi a scala più piccola rispetto a quella desiderata. A parte la semplicità del polinomio interpolante utilizzato, che non consente di estrarre con la stessa precisione di funzioni più complesse (ad esempio quelle a base radiale) le derivate della superficie, il problema maggiore risiede nella forma regolare della maschera di analisi che, essendo quadrata, ingloba celle di elementi non appartenenti alla forma di interesse. Difatto tali celle costituiscono degli *outlier* rispetto alla forma che si vuole analizzare. Risulta chiaro che i bordi delle forme specifiche di interesse non possono essere estratti da un classificatore non supervisionato impiegando il metodo "pixel based". Ciò poiché non sono delineate manualmente come nel caso della classificazione supervisionata.

E' stato quindi affrontato il problema attraverso la ricerca di un criterio che consentisse di aggregare le celle del *raster* in maniera automatica, cioè senza la necessità di delineazione manuale. Tale aggregazione ovviamente è mirata a delimitare aree del DTM secondo criteri geomorfologici: le aree devono

corrispondere il più possibile ad elementi specifici del paesaggio che il geomorfologo è abituato a cartografare attraverso l'analisi tradizionale.

Il metodo dell'analisi a finestre multiple risulterà comunque utile, come vedremo, nel caso della classificazione supervisionata poiché in quel caso la distribuzione dei valori (degli elementi delineati manualmente) viene definita a priori tramite la delineazione di un “training dataset”. Le variabili estratte tramite tali finestre sono equivalenti ad una media pesata dei parametri morfometrici nell'intorno della cella ed utilizzati per il calcolo della distanza statistica dal “training dataset”. I parametri della distribuzione multivariata del “training dataset” sono per definizione corretti. Nel caso del *clustering* invece, non essendovi “training dataset” non vi è una distribuzione multivariata definibile a priori. Un altro motivo per cui è stato scartato il metodo delle finestre multiple è la non convenienza nell'utilizzare l'uso di metodi che impiegano un gran numero di variabili. Gli algoritmi di *clustering* partizionale utilizzati sono di tipo iterativo, per cui il costo in termini computazionali diventa un fattore non trascurabile. La complessità computazionale ad ogni singola iterazione dell'algoritmo k-means è proporzionale al numero di variabili moltiplicato il numero di celle del *raster*. Il metodo sviluppato utilizza principalmente il k-means quale algoritmo di *clustering* (ma può utilizzare indifferentemente anche la versione fuzzy che però è computazionalmente molto più costosa a causa del calcolo della matrice dei valori membro). Poiché la classificazione non supervisionata comporta l'interpretazione e l'assegnazione a posteriori delle classi, la rapidità di esecuzione risulta essenziale poiché il geomorfologo deve poter ottenere più realizzazioni sperimentando diverse soluzioni in termini di numero di classi, al fine di poter scegliere la soluzione considerata maggiormente soddisfacente in base alle forme evidenziate nell'area di studio e alla finalità dell'elaborato.

Al fine si è voluto procedere utilizzandolo come idea di base l'approccio di Speigth ed automatizzarne la metodologia, inglobando comunque nuove idee ed utilizzando parametri molto più accurati e tecniche di raggruppamento di tipo non soggettivo. La fase di sperimentazione ha comportato sia la verifica delle procedure da adottare che dei parametri meglio adatti allo scopo.

Una prima fase è stata quella dell'analisi preliminare dei risultati del clustering delle variabili morfometriche utilizzando i valori estratti dalle singole celle del DTM. Tale metodo non tiene conto dei rapporti tra le variabili nello spazio multivariato e nel reale spazio geografico. Difatti celle adiacenti nello spazio geografico possono essere distanti nello spazio degli attributi più di celle lontane nello spazio geografico se la loro distanza euclidea nel primo, definita come

$$\sqrt{\sum_{l=1}^d (x_l - y_l)^2}$$

dove  $x_l$  ed  $y_l$  sono gli elementi  $l$ -esimo delle variabili  $d$ -variate associate alla cella  $x$  ed  $y$  del raster, è più distante rispetto alla distanza euclidea nello spazio geografico di celle adiacenti.

La distanza tra due celle  $x$  e  $y$  nello spazio geografico è semplicemente

$$\sqrt{(Xcoord_x - Xcoord_y)^2 + (Ycoord_x - Ycoord_y)^2}$$

dove  $Ycoord$  ed  $Xcoord$  sono le coordinate geografiche

Diversi autori (Dikau 1989, Pike 1988), nell'approccio all'analisi della superficie terrestre, considerano di fondamentale importanza la scelta dei parametri, considerando alcuni più utili di altri alla caratterizzazione e distinzione delle varie forme o, considerato l'utilizzo di modelli raster, di singole celle dell'area di studio. Dikau ad esempio considera molto utile per la classificazione del pendio elementi minimali, quali aree omogenee per pendenza ed aspetto. Questi ed altri autori puntano spesso l'attenzione su come sia in linea teorica possibile realizzare una caratterizzazione esaustiva del versante tramite le combinazioni di concavità e convessità nella direzione del profilo e delle curve di livello (curvatura planare). Si è quindi ipotizzato di sfruttare quali principali variabili

morfometriche maggiormente caratterizzanti la superficie terrestre la pendenza, e le due curvature tangenziale e del profilo. Le due curvature utilizzate godono della proprietà di ortogonalità geometrica ma non è detto che siano statisticamente non correlate. Tutte e tre le variabili risultano essere caratterizzate da un coefficiente di correlazione comunque molto basso anche se non trascurabile, soprattutto tra le due curvature (tabella 7.2).

	Pendenza	Curvatura del profilo	Curvatura tangenziale
Pendenza		-0.009372	0.113955
Curvatura del profilo			0.327075

Tabella 7.2 – Correlazione tra le variabili morfometriche utilizzate

D'altronde ciò non ha grande rilievo dal punto di vista della classificazione automatica poiché i coefficienti di correlazione non sono indicativi quanto una covarianza spaziale (Isaaks & Srivastava 1989). Non vi è inoltre relazione tra come la nuvola di punti è distribuita attorno al suo centroide e la posizione delle celle del *raster*. Difatti le coordinate geografiche non sono incluse tra le variabili dei vettori degli attributi. Questo costituisce un limite del *clustering* applicato a dati georeferenziati quando non si utilizza un algoritmo o metodo che tiene conto anche della vicinanza geografica. Infatti, includere i valori delle coordinate geografiche tra gli elementi dei vettori degli attributi, ed effettuare successivamente un *clustering* del set di dati, produce solamente una partizione dello spazio geografico del *raster* con risultati del tutto diversi da quelli desiderati. Tra le ipotesi di lavoro si è anche considerato, in via preliminare, una classificazione basata, oltre che sugli attributi morfometrici, sulla mutua distanza tra le celle. Tuttavia è chiaro che una classificazione di questo tipo necessita del calcolo di una matrice delle distanze che occuperebbe uno spazio di memoria pari a  $N^2/2$  dove N è il numero di celle del raster. Sarebbe proibitivo effettuare il clustering di una tale mole di variabili. Infatti, per tale motivo, le matrici di distanza vengono impiegate solamente per effettuare *clustering* di tipo gerarchico su un numero limitato di valori (Jain et alii 1998).

Esaminando lo *scatter plot* degli attributi pendenza, curvatura del profilo e curvatura tangenziale (fig 7.2) del set di dati originale, è possibile notare la presenza di un unico *cluster* compatto attorno al centroide. In pratica non esistono diversi *cluster* naturali nello spazio degli attributi. Per “cluster naturali” (Webb, 2002) si intendono *cluster* compatti: in pratica due o più nuvole di punti aggregati attorno al loro centroide caratterizzati da aree vuote o meno dense di punti tra una nuvola e l'altra. Questi *cluster* costituiscono eventualmente differenti classi che non si sovrappongono nello spazio multivariato e che sono quindi facilmente separabili da un classificatore. Se i *cluster* hanno forma sferica sono facilmente separabili tramite l'uso della distanza euclidea. Se i cluster risultano separati, nel senso che esistono aree di minore densità tra un cluster e l'altro, ma la distribuzione attorno al loro centroide non è sferica, altri tipi di distanze possono essere impiegati od alternativamente si possono impiegare classificatori di tipo non lineare. Essendo il metodo basato sul k-means, l'uso di classificatori non lineari non può essere impiegato. Inoltre non vi è evidenza di alcun tipo di struttura all'interno del set di dati multivariato, i cui *pattern* appaiono concentrati attorno ad un unico centroide, per cui un *clustering* basato sulla semplice distanza euclidea risulta comunque in questi casi la scelta più appropriata. I vettori delle variabili costituiscano perciò un “continuum” nello spazio multivariato e quindi la distanza euclidea definisce in questo caso più che *cluster* ipersferici, degli “iperpoligoni” del tutto simili ai poligoni di Voronoi. Un'alternativa spesso utilizzata nel caso di classi non separabili linearmente, è quella di effettuare una trasformazione delle variabili di modo da renderle separabili nel nuovo spazio degli attributi (Webb 2002). Si è quindi proceduto alla sperimentazione di trasformazioni del set di dati al fine di ottenere una migliore separabilità, nello spazio multivariato, dei pattern appartenenti alle forme del rilievo di maggior interesse. La separabilità nello spazio multivariato deve in questo caso portare ad ottenere *cluster* che corrispondono nello spazio

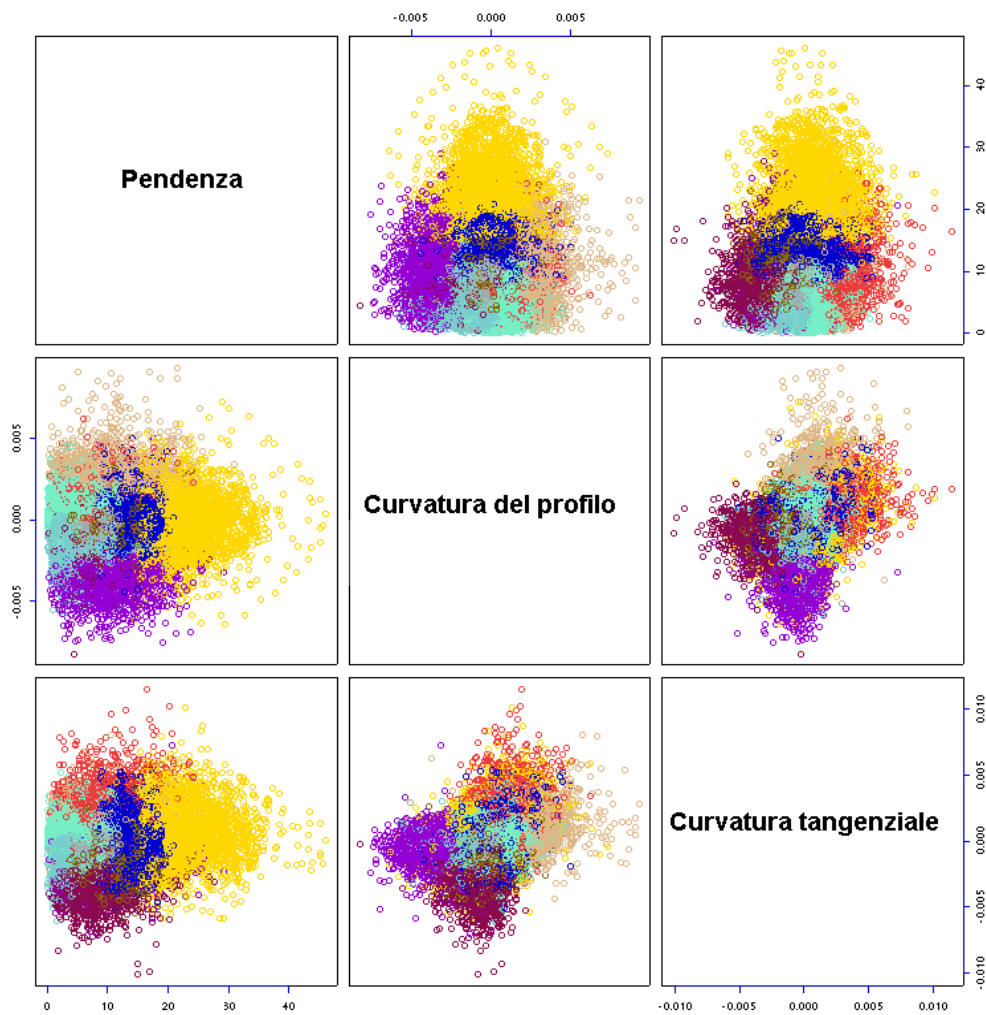


Fig. 7.2 Scatterplot dei parametri morfometrici estratti con il metodo pixel based

geografico, cioè sul raster del DTM, a celle appartenenti alle forme desiderate. Ciò ha presupposto la sperimentazione di diversi tipi di trasformazioni, la successiva separazione in classi a partire dai vari set di dati trasformati, e la validazione del risultato verificando l'effettiva migliore corrispondenza tra classi separabili nello spazio multivariato e classi riconducibili alle forme suddette.



In tal senso la combinazione unica degli attributi costituisce un approccio utile ad effettuare trasformazioni del set di attributi estratti dal DTM. L'elaborazione mirata all'individuazione dei criteri consiste in varie fasi:

- 1) scelta degli attributi morfometrici da utilizzare per ottenere le combinazioni uniche: Il tipo di analisi che si voleva realizzare non si basava direttamente sulla combinazione unica degli attributi riclassificati, ma piuttosto su una analisi statistica basata sui poligoni realizzati, definiti di seguito indifferentemente UCU o “zone”, poiché sono analoghe agli insiemi di celle impiegate nelle analisi zonali dei raster, da una combinazione unica di un sottoinsieme di parametri. Infatti la combinazione unica di 3 o più parametri produce un numero molto elevato di combinazioni e quindi di zone. Si è quindi ipotizzato che utilizzare non più di due parametri considerati maggiormente discriminatori ai fini della classificazione costituissero una scelta adeguata da utilizzare in seguito come base per analisi successive. Si sono quindi utilizzati solo i due parametri caratterizzati dal maggior grado di autocorrelazione spaziale (Isaaks & Srivastava 1989), ossia la pendenza e la quota, riclassificati ognuno in poche classi manualmente. Tale elaborazione è stata realizzata scegliendo esplicitamente il numero di classi tramite il comando GRASS “r.cross” (fig. 7.3). I poligoni (che in realtà corrispondono ad insiemi di celle) risultanti non sono stati utilizzati per raggruppare in classi l'elaborato ma per realizzare una maschera da utilizzare nelle operazioni successive. Tale elaborato costituisce in effetti una segmentazione del raster basata sulle classi di pendenza e quota.
- 2) Una fase di cosiddetto *clumping* ovvero di assegnazione di un record identificativo od *object id*, tramite il comando GRASS r.clump. Tale comando attribuisce un record identificativo differente agli insiemi di celle di zone non adiacenti secondo lo schema di fig. 7.3.
- 3) Una fase di filtraggio dei raster di attributi scelti di seguito, tramite la maschera risultante dalla fase di “clumping”

**r.cross** crea una mappa raster che rappresenta la combinazione unica di tutte le singole categorie presenti nelle mappe in input. E' possibile combinare due o più mappe. Il risultato è una mappa con un numero di classi uguale al prodotto del numero di classi delle mappe in input. La mappa finale può essere utilizzata per ricercare i clump attraverso il comando **r.clump**.



mappa 1 categoria	mappa 2 categoria	mappa finale categoria
1	1	1
2	2	2
1	2	3
2	1	4

**r.clump** cerca tutte le aree formate da celle contigue assegnate alla stessa categoria in una mappa raster in input. Come output assegna un'unica categoria alle celle così individuate. La distinzione tra le celle con categorie diverse nella mappa iniziale viene conservata, ciò significa che se celle con categorie differenti sono adiacenti esse non saranno aggregate nella stessa categoria. L'algoritmo utilizza un filtro 2x2 che raggruppa le celle secondo il criterio descritto solo per colonne e per righe, in questo modo le celle lungo la diagonale non vengono considerate contigue (vedi illustrazione).

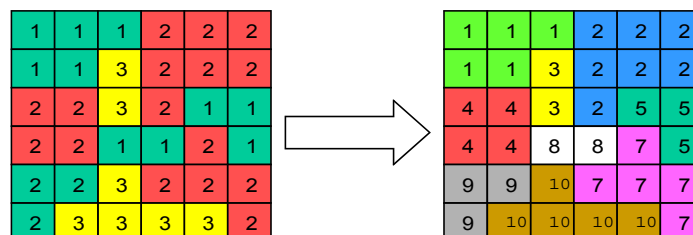


Fig 7.3 Logica di funzionamento dei comandi **r.cross** ed **r.clump** del pacchetto **gis GRASS**.

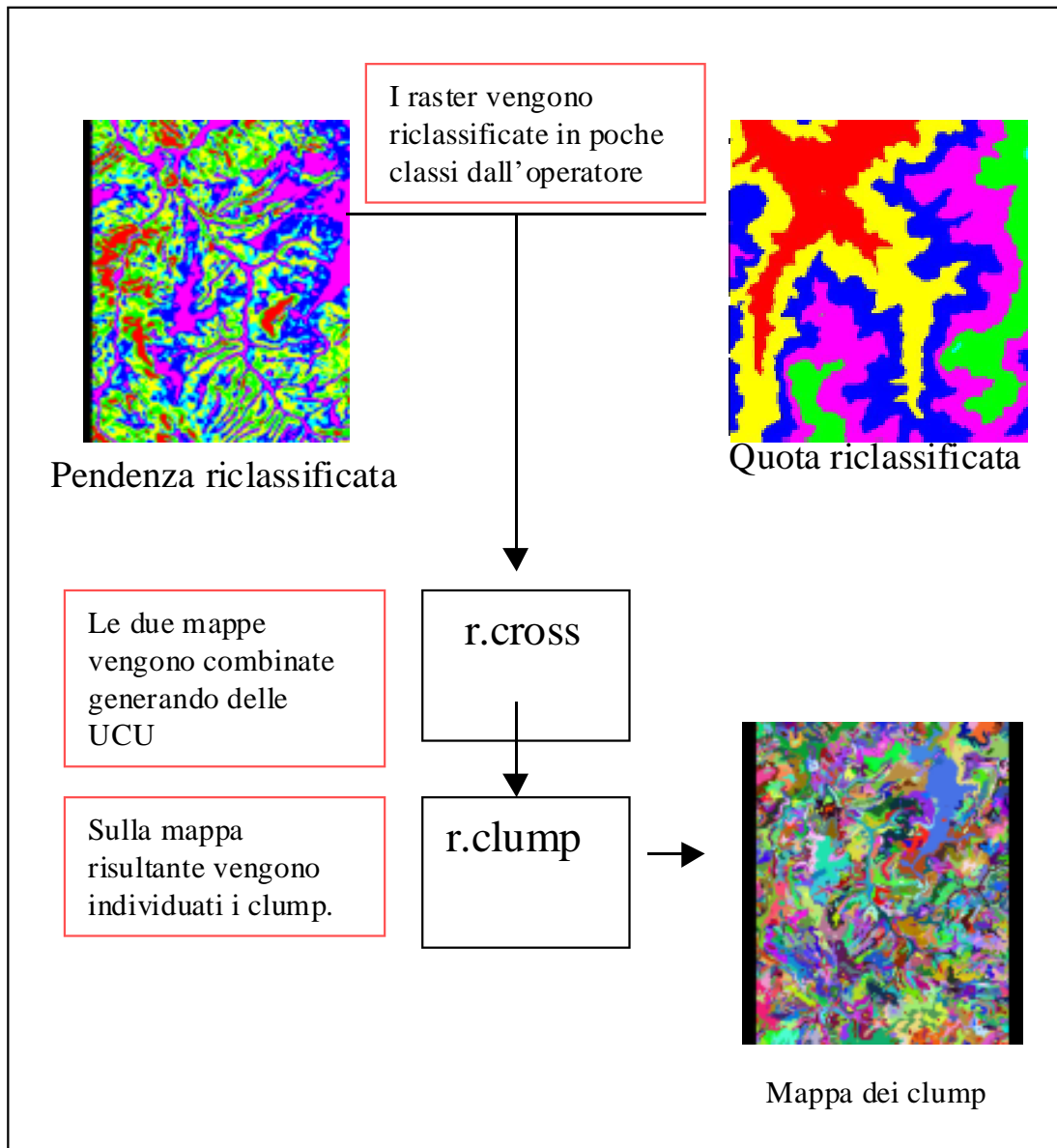
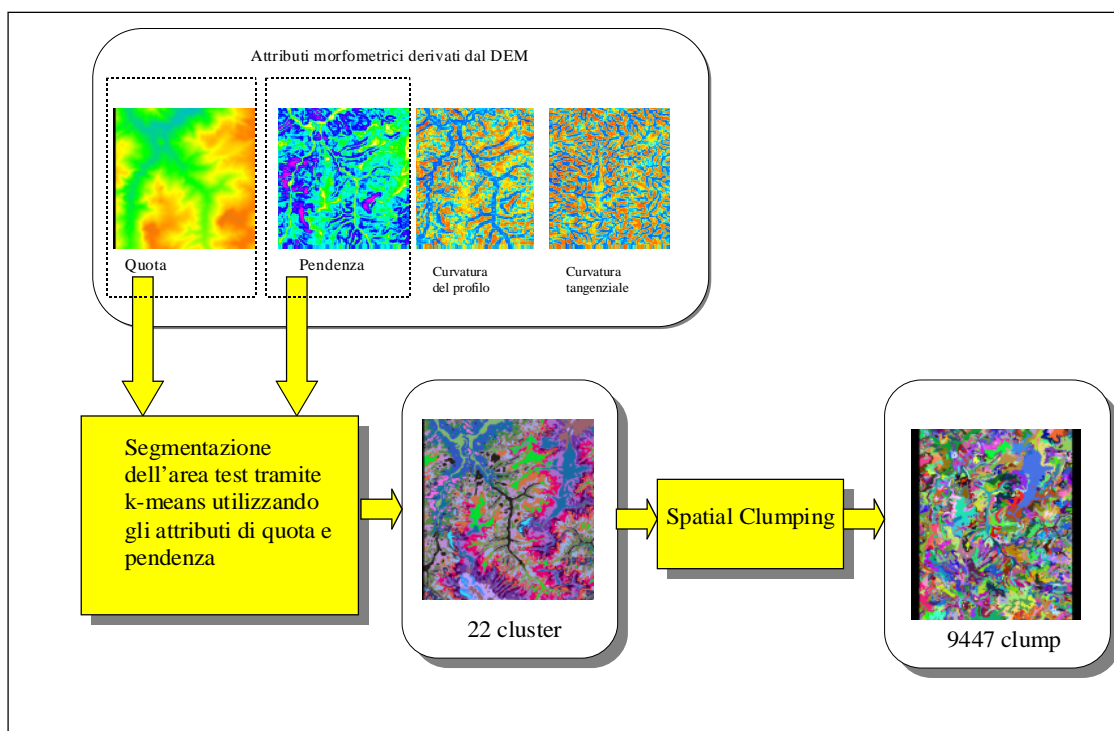
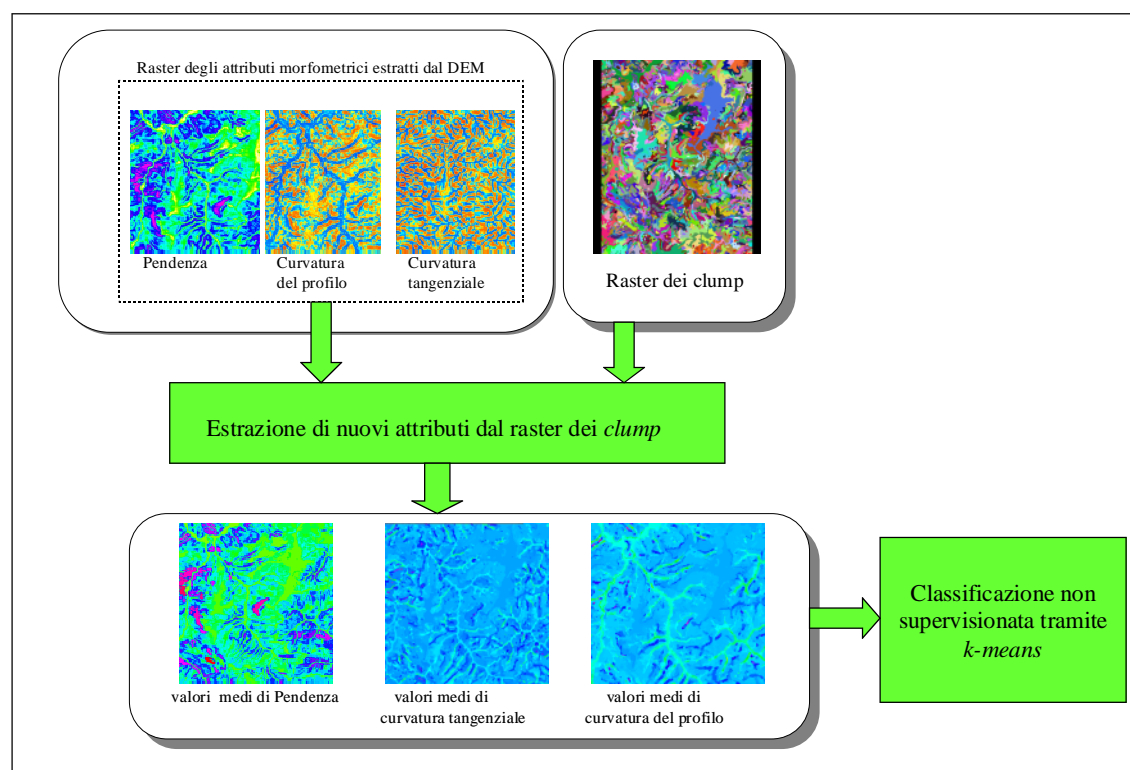


Fig 7.4 Flusso del procedimento di generazione dei clump basato sulla combinazione unica di raster riclassificati manualmente



A)



B)

Fig 7.5 – Diagramma di flusso del metodo di classificazione non supervisionata proposto

A) Generazione dei clump mediante k-means B) Estrazione dei nuovi attributi dai clump e clustering

La fase di “clumping” restituisce un nuovo raster i cui elementi minimi (i clump) possono essere caratterizzati tramite i parametri morfometrici estratti dalle celle che ricadono all'interno degli stesi. In pratica diventa un raster di copertura per eventuali analisi di altri livelli informativi.

Per la fase di generazione delle UCU, sono state sperimentate varie soluzioni in termini di parametri morfometrici da combinare. Innanzitutto i raster dei parametri morfometrici utilizzati di volta in volta sono stati riclassificati in un numero ridotto di classi per evitare la generazione di un numero di combinazioni uniche eccessivo. Nel caso delle due curvature ad esempio, data l'elevata caoticità della distribuzione spaziale di tali attributi, riconducibile ad un basso grado di autocorrelazione spaziale, ci si è limitati a riclassificare i raster nelle due classi concavo e convesso, ovvero attribuendo il valore 0 alle curvature con segno positivo ed 1 a quelle con segno negativo. Ciò considerando che la differenza nella forma di elementi basilari, quali creste od impluvi, è caratterizzata proprio da concavità o convessità nella curvatura tangenziale. Anche l'analisi del profilo si basa sulla suddivisione dello stesso in aree concave e convesse, cioè matematicamente sul segno della curvatura del profilo (Young, 1972). Va considerato anche che tale fase è mirata solo ad una suddivisione preliminare e non della classificazione vera e propria. Quota e pendenza, data la maggiore autocorrelazione spaziale, sono stati riclassificati in più classi. Dopo aver verificato visivamente i differenti risultati, si è deciso di utilizzare questi due ultimi parametri e scartare le altre soluzioni. Dal punto di vista teorico, poiché questi due parametri rappresentano rispettivamente la superficie topografica e la sua derivata prima, la loro combinazione comporta la suddivisione del versante lungo le principali rotture di pendenza. Questi attributi, che sono più immediatamente legati alla geometria topografica, sono evidentemente più efficaci degli altri attributi utilizzati nelle prove, i quali sono correlati alla forma della superficie da leggi che forse sono meno legate ai processi genetici delle superfici di interesse e che sono certamente meno coincidenti con quelle che

L'operatore geomorfologo è abituato ad utilizzare (e a ritenere significative) quando scompone un paesaggio in forme composite. Da diversi tentativi si è notato in particolare che l'utilizzo in questa fase di concavità e convessità non evidenzia elementi normalmente percepiti come significativi dal punto di vista geomorfologico, similmente a quanto accade nel caso del clustering basato sulla analisi degli attributi associati alle singole celle. L'ipotesi è ovviamente che proprio questi stessi parametri presi in considerazione, nonostante ridotti in poche classi, non evidenzino altro che bordi associati ad elementi troppo specifici (quali appunto impluvi e displuvi). Inoltre si può facilmente dedurre come i clump definiti dai bordi di aree concave [convesse], racchiudano al loro interno solo elementi concavi [convessi] i quali, durante la fase di *clustering* basata sulla suddivisione in poche classi, saranno raggruppati nella medesima classe, mentre le forme composite sono caratterizzate da insiemi di elementi con differente grado di concavità e convessità sia tangenziale che del profilo, oltre che differente angolo di pendenza, lungo il versante. Per tale motivo i valori di concavità e convessità sono stati impiegati solo successivamente per caratterizzare i singoli clump per realizzare una classificazione finale in forme come intese sopra.

La successiva fase di analisi comporta l'estrazione di nuovi attributi all'interno delle aree di clump, impiegando ad esempio il comando "r.statistics" di GRASS. Questo è un comando che serve a generare un'analisi zonale, estraendo diversi valori statistici ed utilizzando come maschera un raster di copertura, ad esempio anche il raster dei clump. A tal fine, essendo i raster impiegati a doppia precisione, ovvero con precisione di 64 bit, non è stato possibile utilizzare tale comando poiché si basa su un algoritmo che può elaborare solo valori numerici interi ossia di tipo *integer*. I raster utilizzati in GRASS possono contenere più campi oltre al valore numerico. Per poter elaborare valori in virgola mobile, tale comando, deve utilizzare il campo "categoria" (che corrisponde all'etichetta della cella) il quale non è un valore numerico ma un dato di tipo alfanumerico e che

deve essere convertito in valore numerico dall'algoritmo in un processo che risulta estremamente lento. A tal fine è stato necessario scrivere un algoritmo (denominato *r.fpstats*) ed integrarlo come modulo di GRASS, in grado di estrarre direttamente dal valore numerico in formato 64 bit, i valori statistici indicati in tab 7.2.

Dei quattro raster utilizzati per caratterizzare il DTM solo tre sono stati utilizzati in questa fase. Attraverso numerose prove, sono estratti diversi descrittori statistici e valutati visivamente i risultati ottenuti scegliendo differenti numeri di cluster e differenti attributi estratti dai clump. Sono stati così individuati gli attributi più idonei sia per la fase combinatoria che i descrittori statistici per la fase di classificazione finale in elementi del paesaggio tramite *clustering*. Questa fase di sperimentazione era mirata ad individuare criteri validi dal punto di vista generale, cioè tarare, in maniera empirica, per tentativi e valutazioni visive, il metodo da impiegare e che verrà successivamente validato comparandolo con una classificazione “pixel based”. Il numero di classi da utilizzare, una volta definito il metodo, rimane comunque dipendente dall'area di studio e le classi finali ottenute devono essere comunque oggetto di valutazione (ed ovviamente interpretazione) come si vedrà per le aree analizzate. In tabella 7.2 sono evidenziati in grassetto i parametri e attributi scelti per il metodo e le altre combinazioni sperimentate.

<i>Raster degli attributi morfometrici</i>	<i>Nuovi attributi estratti dai clump</i>				
Quota	Deviazione standard	Range			
<b>Pendenza</b>	<b>Media</b>	Deviazione standard	Massimo	Minimo	Range
<b>Curvatura del profilo</b>	<b>Media</b>	Deviazione standard	Massimo	Minimo	Range
<b>Curvatura tangenziale</b>	<b>Media</b>	Deviazione standard	Massimo	Minimo	Range

*Tabella 7.2*

### ***Segmentazione basata sull'analisi bivariata***

Il metodo basata sulla combinazione unica è stato successivamente sostituito da una tecnica più avanzata per produrre le zone e quindi i clump (fig 7.5). I bordi dei clump ottenuti a partire dalle UCU sono il risultato dell'intersezione tra le aree prodotte dalla riclassificazione manuale dei due attributi utilizzati. Ciò produce, come spesso in tale tipo di analisi, un gran numero di aree di piccola estensione e bordi che non rispecchiano in modo preciso il naturale andamento delle rotture di pendenza perché influenzati dalla forma dei bordi delle classi di pendenza e quota. In particolare il raster della quota riclassificata rispecchia l'andamento delle isoipse. Si è quindi sperimentato un metodo basato sulla analisi bivariata di quota e pendenza. A tal fine si è utilizzato il k-means. In tal modo non risulta più necessario riclassificare i due raster manualmente ma solamente indicare il numero di classi che si intende ottenere dalla combinazione dei due parametri. Questo metodo è caratterizzato da un doppio vantaggio. Il metodo diventa meno soggettivo riducendo ad una sola la scelta del numero di classi in cui suddividere il DTM al fine della realizzazione dei clump. Si realizzano clump dai bordi più naturali, essendo questo il prodotto della analisi congiunta della varianza dei due parametri. Come visto nel paragrafo 5 infatti il k-means (o anche la versione fuzzy) producono classi minimizzando la varianza interna dei cluster. Al fine di automatizzare l'intera procedura sono stati realizzati degli script della Shell di Unix che utilizzano i comandi `r.clump`, `r.fpstats`, `r.clustering` ed `r.fuzzy`. (in appendice).



Pendenza	Curvatura del profilo	Curvatura tangenziale
Min. : 0.0432	Min. :-8.194e-03	Min. :-1.014e-02
1st Qu.: 7.5931	1st Qu.: -5.491e-04	1st Qu.: -7.295e-04
Median :10.5174	Median : 1.714e-05	Median : 7.660e-05
Mean :10.9242	Mean :-6.156e-05	Mean : 5.845e-05
3rd Qu.:13.7578	3rd Qu.: 5.313e-04	3rd Qu.: 8.622e-04
Max. :46.0849	Max. : 9.260e-03	Max. : 1.150e-02

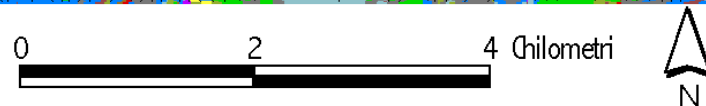
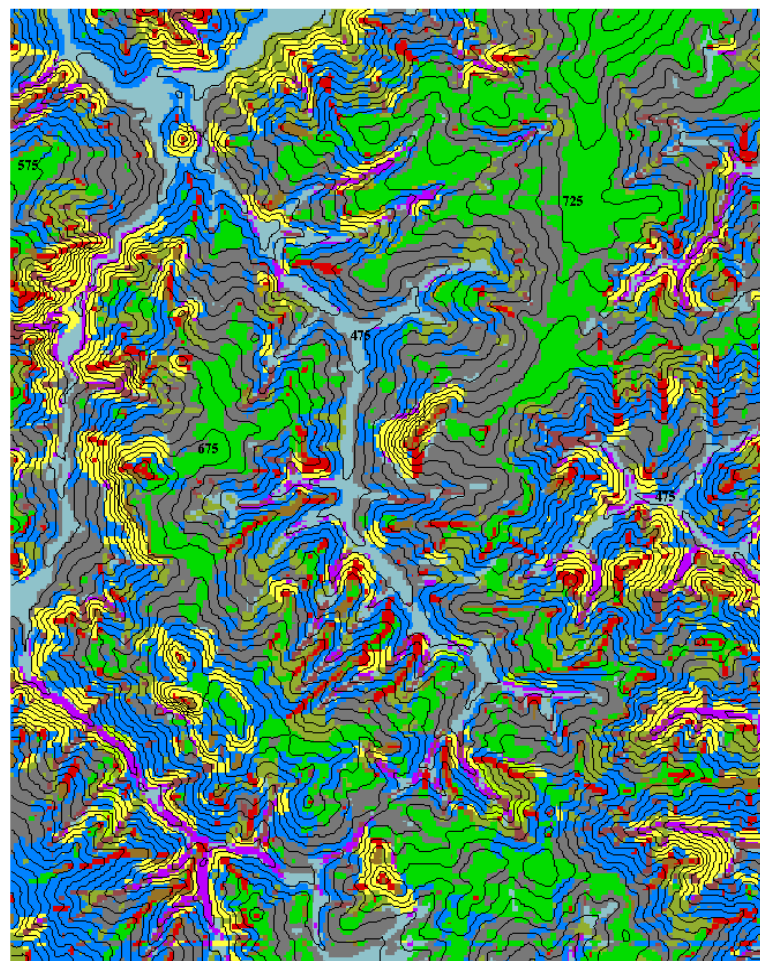
*Tabella 7.3a - Sintesi statistiche degli attributi utilizzati nel metodo "pixel basedl"*











Pendenza	Curvatura del profilo	Curvatura tangenziale
Min. : 0.2431	Min. :-7.364e-03	Min. :-9.903e-03
1st Qu.: 7.2778	1st Qu.: -1.932e-04	1st Qu.: -2.173e-04
Median :10.1473	Median : 3.599e-05	Median : 4.719e-05
Mean :10.9242	Mean :-6.156e-05	Mean : 5.845e-05
3rd Qu.:13.2238	3rd Qu.: 2.138e-04	3rd Qu.: 3.962e-04
Max. :34.3980	Max. : 9.260e-03	Max. : 1.150e-02

*Tabella 7.3b - Sintesi statistiche degli attributi utilizzati nel metodo proposto*

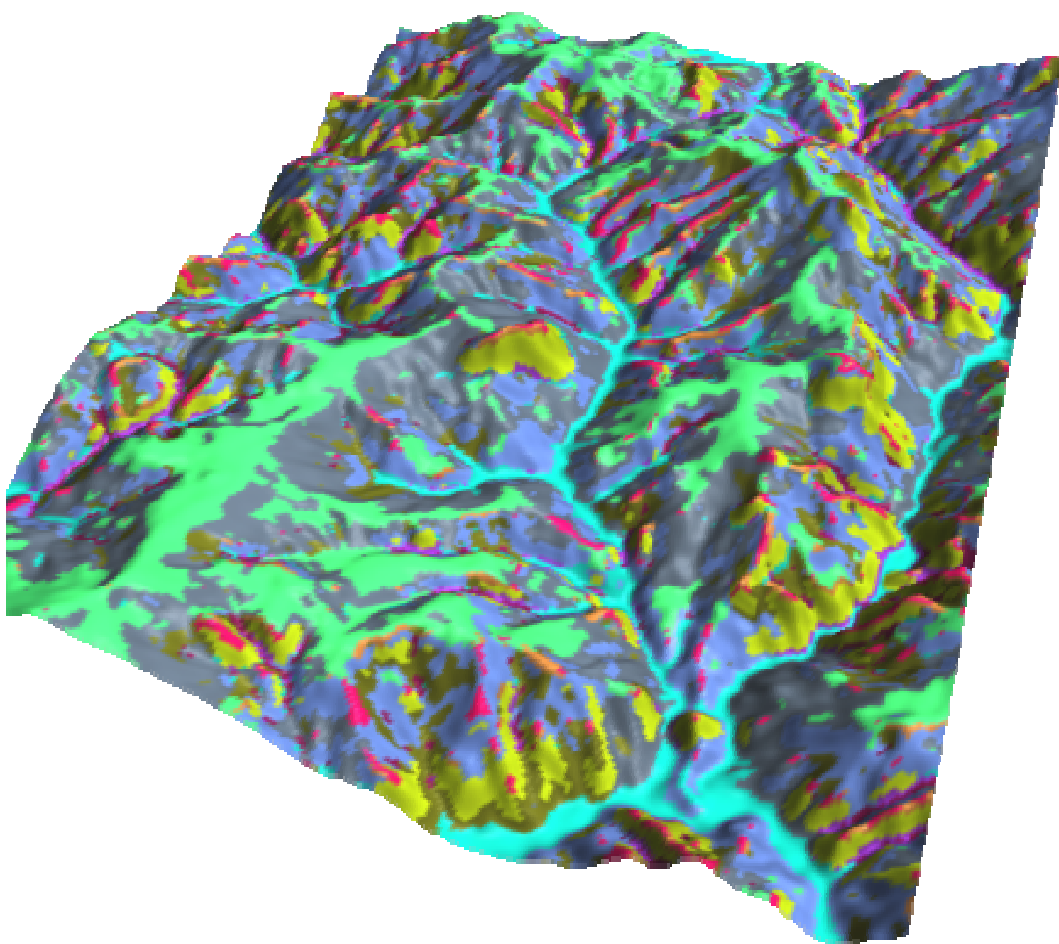
Cluster	numero celle	pendenza	curvatura del profilo	curvatura tangenziale
1	5626	-937.29.00	-892.28.00	-823.09.00
2	58971	36.59.00	-508.16.00	200.48.00
3	3096	441.32.00	318.00.00	398.32.00
4	11430	-348.27.00	951.00.00	773.35.00
5	19161	-554.38.00	269.00.00	698.22.00
6	45040	-02.30.00	370.33.00	-33.05.00
7	53205	-22.35.00	-951.14.00	-122.57.00
8	34884	77.36.00	-288.09.00	303.31.00
9	16994	826.45.00	-538.45.00	-39.32.00
10	17412	-183.18.00	699.34.00	117.56.00
		pendenza	curvatura del profilo	curvatura tangenziale
1	10032	-784.23.00	673.10.00	927.05.00
2	38125	-367.37.00	719.50.00	785.02.00
3	72616	-530.31.00	301.00.00	-652.51.00
4	72400	-615.59.00	-184.01.00	-550.31.00
5	1514	-829.08.00	-789.14.00	-895.22.00
6	24405	526.24.00	395.44.00	939.39.00
7	55727	668.20.00	-1023.09.00	-350.02.00
8	79718	-422.23.00	-996.51.00	-1069.47.00
9	67652	701.29.00	431.58.00	750.41.00
10	6402	-773.50.00	954.38.00	825.09.00

*Tabella 7.4 Centroidi risultanti dal clustering "pixel based" in alto, dal clustering mediante il metodo proposto in basso. Gli attributi del secondo risultato derivano dalla media sui clump.*



- |                                                                                     |                                              |                                                                                     |                                                        |
|-------------------------------------------------------------------------------------|----------------------------------------------|-------------------------------------------------------------------------------------|--------------------------------------------------------|
|  | 1 - Crinali arrotondati                      |  | 6 - Crinali affilati                                   |
|  | 2 - Fondovalle a V                           |  | 7 - Testate vallive acclivi                            |
|  | 3 - Glacis e superfici di erosione sommitale |  | 8 - Fondo di valleciole di basso ordine gerarchico     |
|  | 4 - Fondo di valli a V aggragate             |  | 9 - Elementi a controllo strutturale o sub-strutturale |
|  | 5 - Versanti fluvio-denidazionali            |  | 10- Fianchi vallivi acclivi ed elementi di interfluvio |

*Fig 7.6a Risultato della classificazione del raster C tramite classificazione non supervisionata utilizzando il metodo proposto*



*Fig 7.6b Vista 3D del raster C classificato tramite il metodo proposto*

### ***Validazione numerica***

Oltre alla analisi visiva ed interpretazione delle classi ottenute, è stata fatta una comparazione analitica tra il risultato del clustering “pixel based” ed il risultato ottenuto con il metodo proposto. Come descritto nel capitolo 5, il fuzzy c-means calcola una matrice membro e quindi ad ogni cella del raster è associato un vettore di valori membro, i cui elementi indicano il grado di appartenenza a quella classe. I valori vanno da 0 ad 1 e la loro somma è 1. In pratica nel nostro caso vi sono 10 valori membro. Un valore pari a zero indica che quella cella non ha nessuna affinità con quella classe, mentre un valore pari ad uno indica che la cella appartiene sicuramente a quella classe. Si tratta ovviamente di valori limite ed il fuzzy c-means produce sempre valori intermedi. Le celle vengono quindi classificate in cosiddette *hard classes* cioè assegnate univocamente alla classe con il valore membro più grande.

Per insiemi di celle che hanno un valore membro in una classe molto vicino ad uno, i valori per le altre classi, dovendo sommare il loro insieme ad uno, saranno molto bassi e la visualizzazione del raster darà una risultato visivo molto simile alla corrispondente mappa delle *hard classes*. Per celle con valori membro tutti molto simili, non esiste una classe “dominante” e quindi il raster dei valori membro risulterà meno interpretabile o come si dice in gergo “più confuso”. In effetti esiste un metodo per calcolare il grado di confusione associato ad una classificazione fuzzy. Si usa spesso il rapporto tra il secondo valore membro più grande e quello più grande, ossia (Burrough et alii 1997)

$$CI = \frac{\sum_{n=1}^C \text{Maxfuz}_2(n)}{\sum_{n=1}^C \text{Maxfuz}_1(n)} \quad (7.1)$$

dove Maxfuz1 è il valore membro più grande associato alla cella n, Maxfuz2 è il secondo valore membro più grande e C è la cardinalità del *cluster*.

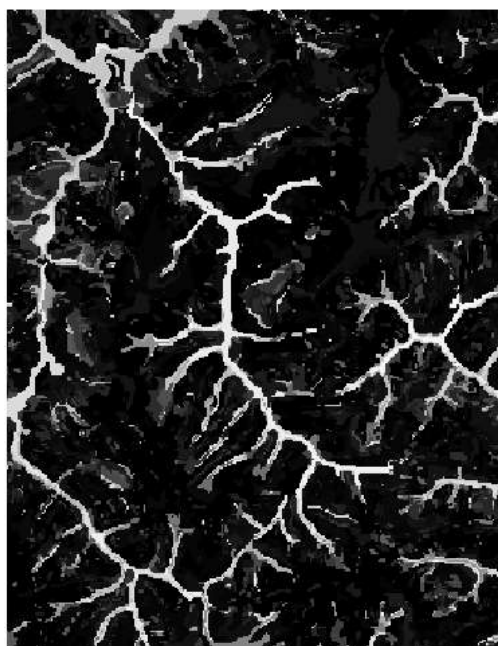
In entrambi i metodi di *clustering* gli attributi morfometrici utilizzati ed il numero

di classi scelte (10) è lo stesso. Si è quindi effettuata una classificazione impiegando come algoritmo di clustering il fuzzy c-means, ma nel caso “pixel based” si utilizzano i tre raster della pendenza, curvatura tangenziale e curvatura del profilo senza ulteriori passaggi, ovvero utilizzando gli attributi associati alle singole celle. Per il metodo proposto sono stati utilizzati i tre raster ottenuti dal calcolo della media sui clump (i raster indicati come nuovi attributi estratti nella fase di figura 7.5b. In tabella 7.4 sono indicati i centroidi delle 10 classi risultanti per i due set di dati. In entrambi i casi la matrice degli attributi è stata riscalata utilizzando il metodo dello scarto medio assoluto e quindi i valori dei centroidi sono valori riscalati e non i reali valori di pendenza e curvatura. Per ogni classe è stato estratto il raster dei valori membro in modo da poter visualizzare la dispersione di tale valore rispetto alle classi interpretate. In figura 7.7 si può notare come la dispersione dei valori membro delle classi per il metodo adottato, sia molto simile al pattern spaziale definito dall'assegnazione delle “hard classes” soprattutto per le classi interpretate come fondovalle e glacis. Nel caso della classificazione “pixel based”, si nota una maggiore dispersione dei valori membro e una generale caoticità nella distribuzione. In tabella 7.5 sono indicate per ogni classe i valori di “Confusion Index”, come si può notare risulta più basso in tutte le classi ottenute con il metodo proposto.

Cluster	Metodi sviluppato CI	Metodo “pixel based” CI
1	0,26	0,47
2	0,33	0,37
3	0,32	0,44
4	0,26	0,33
5	0,39	0,4
6	0,28	0,47
7	0,31	0,44
8	0,34	0,47
9	0,45	0,48
10	0,36	0,43

*Tabella 7.5*

Metodo proposto



Fondovalle CI=0,31

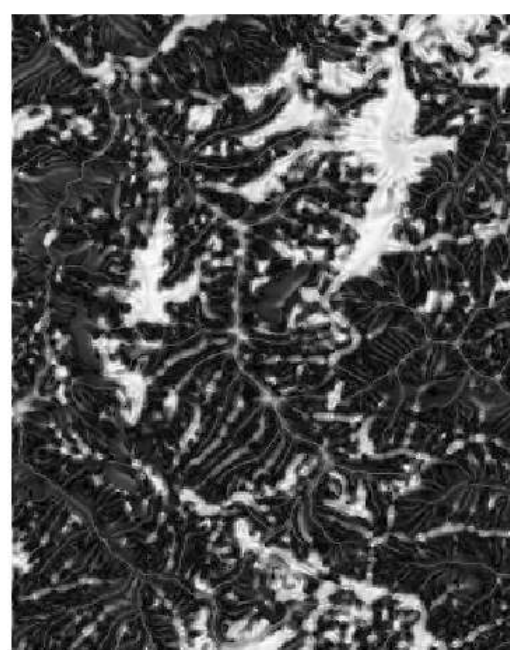
Metodo "Pixel Based"



Fondovalle CI=0,33



Glacis CI=0,26



Glacis CI=0,37

*Fig. 7.7 Raster dei valori membro per le due classi glacis e fondovalle e relativo confusion index estratti dall'area test ricoperta dal raster C.*

Il CI calcolato dalla formula (7.1) tiene conto di tutte le celle appartenenti ad una classe. La confusione elevata significa che una classe, che dovrebbe corrispondere ad un elemento geomorfologico, viene mappato in maniera incerta. Essendo l'indice calcolato come somma delle singole "confusioni", ciò presuppone che in differenti aree del DTM, lo stesso elemento può coincidere con una corretta classificazione mentre in altre aree, dove la confusione è più elevata, può risultare in una classificazione errata. Ciò ovviamente va confrontato anche con analisi visiva delle *hard classes* (fig 7.9) per capire in quali aree effettivamente un metodo può generare una classificazione corretta mentre l'altro erronea.

Perché il metodo produca una classificazione di qualità superiore è necessario tenere conto della aggregazione nello spazio geografico di celle con caratteristiche simili, anche se il *clustering*, avviene nello spazio degli attributi. A tal fine il metodo risulta migliore se nello spazio degli attributi le classi risultano separate, ovvero corrispondono a cluster compatti che non si compenetrano.

Un secondo metodo di verifica numerica della migliore qualità della classificazione è il calcolo dell'indice di dispersione quadratica. Questo consiste nel verificare in maniera oggettiva per due classificazioni differenti che utilizzano set di dati differenti ed un numero di classi identico, quale comporti la formazione di cluster più compatti. Un *cluster* è considerato più compatto di un'altro se la somma degli scarti quadratici è inferiore. Sono stati calcolati due tipi di somme di scarti quadratici dal risultato del *clustering*. La somma degli scarti quadratici entro i cluster (WSSE) e quella totale (TSSE). La prima è la somma degli scarti quadratici tra i vettori associati ad una classe ed il suo centroide. La seconda è la somma totale degli scarti, ovvero la somma degli scarti di tutti i vettori di tutte le classi dal centroide dell'intero set di dati. Questo è come considerare il risultato della classificazione in un'unico cluster, soluzione che consente di comparare il risultato della classificazione per dati distribuiti diversamente nello spazio multivariato. L'indice utilizzato è dunque il seguente

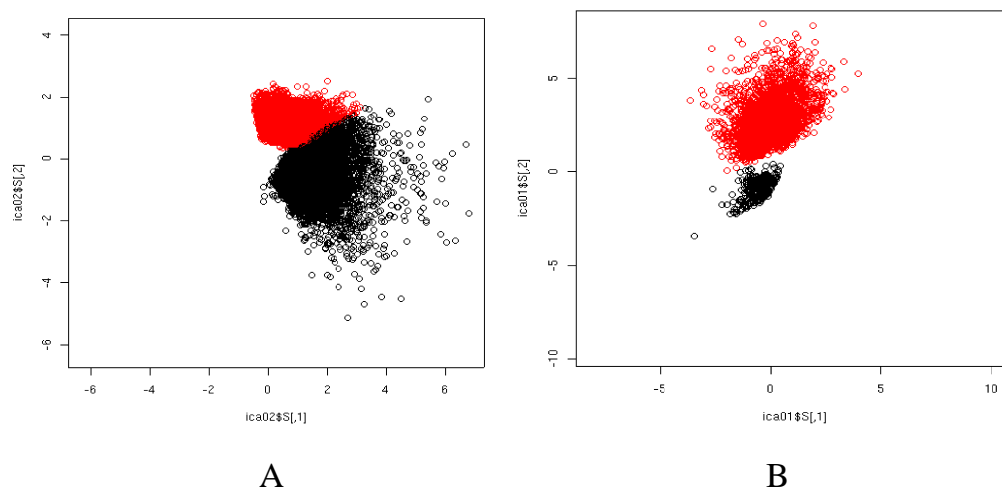


$$1 - \frac{WSSE}{TSSE} \quad (7.2)$$

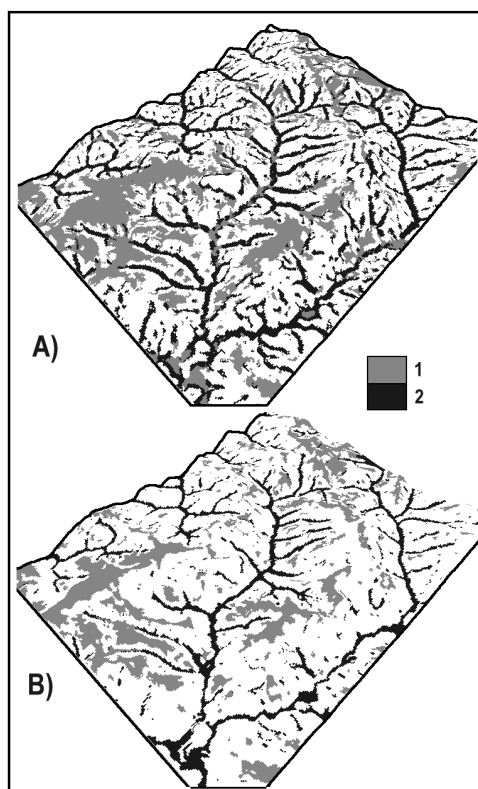
più l'indice è vicino ad uno, migliore si può considerare la compattezza dei cluster nello spazio multivariato. Per la classificazione con il metodo proposto si è ottenuto un valore di 0,915 contro 0,909 del metodo “pixel based”.

Un metodo più diretto e visivamente interpretabile è quello di generare, a partire dalle nuvole di punti associate alle classi di interesse, degli *scatterplot* e verificare il grado di compenetrazione di tali nuvole. Ovviamente essendo lo spazio multivariato a più dimensioni, risulta impossibile utilizzando metodi di visualizzazione ordinari per verificare tale ipotesi; ciò poiché i punti si proiettano su un gran numero di piani e risulta difficile poter visualizzare tale caratteristica oltre le tre dimensioni. Nel caso in esame si sono utilizzate solo tre variabili. Si tratta di spazio a tre dimensioni per cui tramite rotazione degli assi con opportuni software di visualizzazione quali ad esempio Xgobi (Swayne et alii 1998) è possibile, ruotando gli assi delle variabili, trovare un piano di proiezione che consente di verificare o meno tale ipotesi. Tuttavia essendo il metodo di applicabilità generale ed eventualmente utilizzabile con più di tre variabili, risulta più opportuno utilizzare tecniche di “projection pursuit” (Friedman 1987). Tali tecniche consentono di individuare, in uno spazio a più di tre dimensioni, e quindi non facilmente rappresentabile in uno spazio a due, dei piani di proiezione in grado di rivelare strutture interessanti ai fini della classificazione. Nel nostro caso si tratta di identificare un piano di proiezione che evidenzi la separabilità o meno delle classi. Nel caso di uno spazio con tre variabili consente di trovare tale proiezione in maniera automatica. A tal fine, tra tutte le tecniche possibili, si è scelto di utilizzare la analisi delle componenti indipendenti o ICA (Independent Component Analysis) (Lee 1998). Questa è una tecnica che, benché non si stata appositamente sviluppata per applicazioni di “projection pursuit” risulta in effetti adattissima allo scopo. L'ICA consiste infatti in una trasformazione delle componenti principali o PCA (Webb 2002) mirata ad individuare quelle rotazioni

di tali componenti che minimizzano la “gaussianità” della distribuzione multivariata del set di dati. Mentre la PCA identifica le proiezioni lungo le direzioni di massima correlazione, e quindi realizza di fatto una decorrelazione del set di dati, l'ICA cerca le direzioni che annullano i momenti statistici di ordine elevato. Essendo i *cluster* compatti e separati, l'insieme di punti proiettato lungo un asse o un piano che ne enfatizzi la separabilità sarà caratterizzato da una distribuzione meno gaussiana di un unico *cluster* compatto. Ciò si rileva molto utile quando la dimensionalità del set di dati è molto elevato. Si è quindi proceduto selezionando le sole nuvole di punti appartenenti ad alcune classi di cui si era riscontrata la scorretta classificazione da parte del metodo “pixel based” assegnando colori differenti per i *cluster* delle due classi comparate. Per effettuare la trasformazione ICA si è utilizzato il pacchetto “fastICA” (Marchini et alii) in ambiente statistico R (Bivand 2000,1999,1998). Questa particolare implementazione utilizza un metodo basato sulla minimizzazione della “negentropia”, considerato più robusto rispetto alla minimizzazione dei momenti statistici. Lo scatterplot di figura 7.8a mostra la nuvola di punti associata con le celle del DTM estratte dalle classi ottenute con il metodo proposto. Poiché la classificazione ottenuta con tale metodo è stata considerata molto valida, soprattutto nel separare le superfici di interesse dal fondovalle, è possibile paragonare gli insiemi di celle associate ad ogni classe come se fossero state delineate manualmente. Lo scatterplot di figura 7.8a visualizza i valori (nello spazio multivariato) dei parametri morfometrici estratti cella per cella ma ricadenti all'interno delle classi ottenute con il metodo proposto. La proiezione ICA evidenzia come, utilizzando gli stessi parametri morfometrici, il metodo “pixel based” non consente di ottenere classi separabili, poiché è evidente la forte compenetrazione delle nuvole di punti. In figura 7.8b, a puro scopo comparativo, le stesse classi risultano in nuvole di punti isolate se si utilizzano gli stessi attributi utilizzati nel metodo proposto.



*Fig 7.8 Scatterplot prodotti con il metodo ICA, delle classi corrispondenti alle classi “glacis+paleosuperfici” (nero) e “fondovalle+rami di basso ordine+valli a V” (rosso). A) Metodo pixel based B) Metodo proposto.*

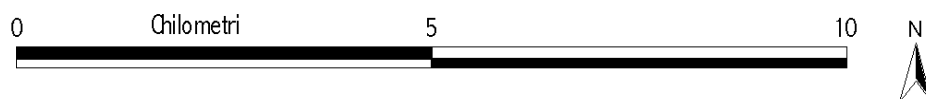
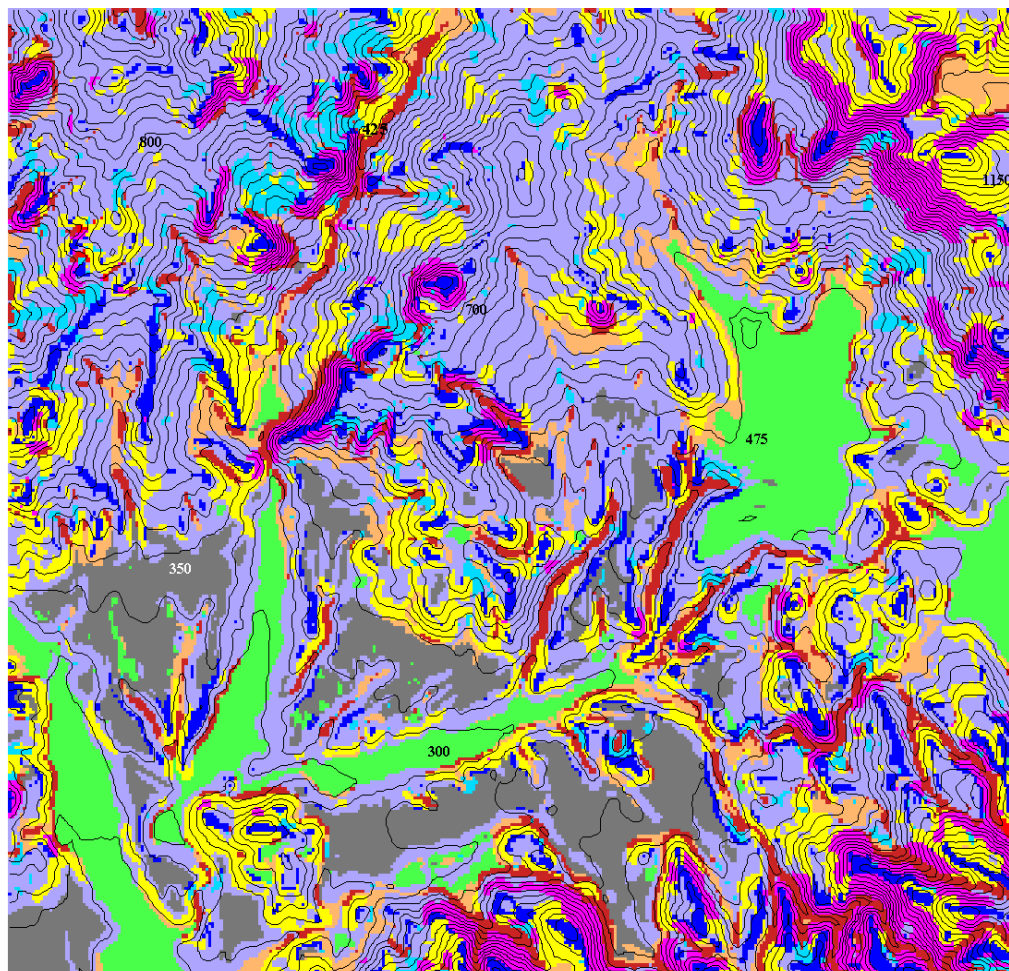


*Fig 7.9 Block diagram del raster C con sovrimposte le classi:*

*1) “glacis+paleosuperfici” e  
2) “fondovalle+rami di basso  
ordine+valli a V”. A) metodo  
“pixel based”*

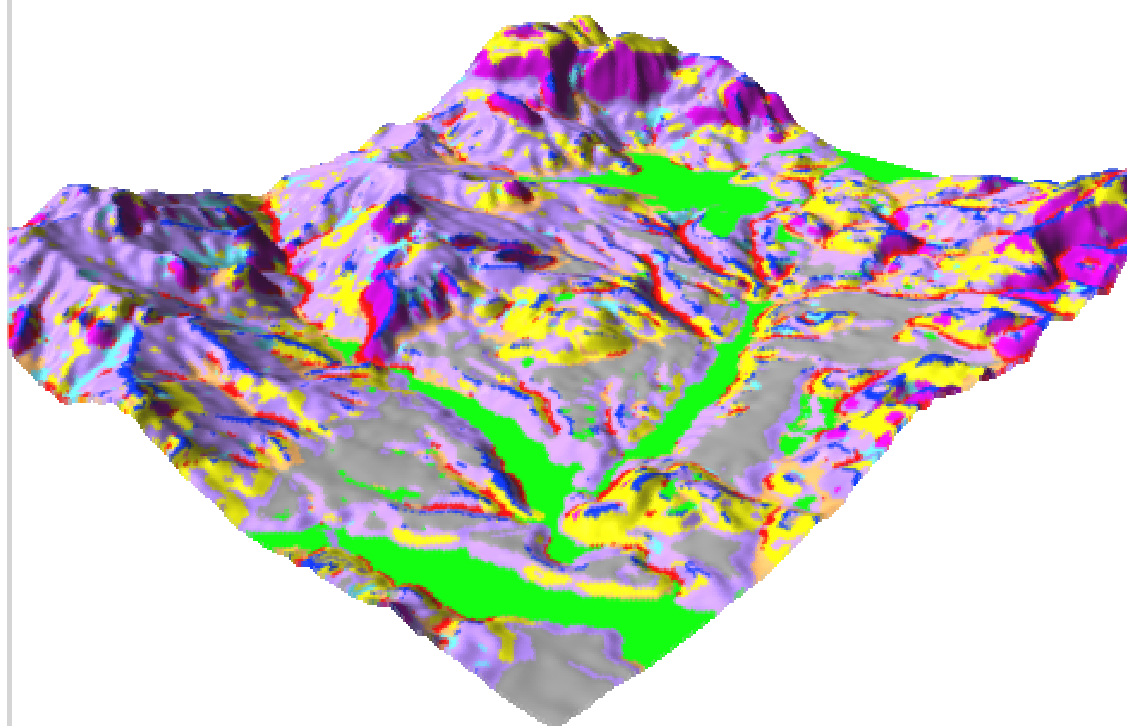
*B) metodo proposto.*

*(da Aucelli et alii 2005)*



- 1 - Aree a debole concavità di diversa origine
- 2 - Crinali
- 3 - Aree di incisione fluviale/Valli a V
- 4 - Elementi a controllo strutturale o substrutturale
- 5 - Scarpate fluviali ed elementi di interfluvio
- 6 - Versante di origine fluvio-denudazionale
- 7 - Fondovalle / Piana alluvionale
- 8 - Terrazzi fluviali
- 9 - Testate vallive a forte acclività / Nicchie di distacco di frane

*Fig 7.10a - Classificazione non supervisionata della piana di Isernia (raster B) mediante il metodo proposto*



*Fig 7.10b Vista 3D del raster della piana di isernia (raster B) classificato mediante il metodo proposto*

In fig. 7.10a ed 7.10b sono evidenziati i risultati (9 classi) per l'area test del raster B. Il metodo, anche in questo caso, si rivela utile a differenziare aree con valori simili o sovrapposti di singoli parametri morfometrici. In questo caso è stato possibile discriminare il fondovalle dai terrazzi di III e IV ordine (accorpati). Si deve solo alla scarsa risoluzione del DTM (40 metri) il fatto che non si sia riusciti ad evidenziare gli altri ordini di terrazzi all'interno dell'area identificata come piana alluvionale (poco ampi e di difficile individuazione anche tramite l'esame delle carte topografiche in scala 1:25.000). Tra gli altri elementi, risultano messi in evidenza in maniera soddisfacente le forme del rilievo a controllo strutturale, i crinali ed elementi che è possibile identificare come nicchie di distacco di frane.

### ***Discussione***

Il metodo proposto utilizza una fase di estrazione di nuovi attributi e li sottopone ad un algoritmo di *clustering* iterativo. Si è evidenziato come sia possibile separare nello spazio degli attributi *cluster* che corrispondono nello spazio geografico ad aggregati di pixel riconducibili più facilmente a forme del paesaggio che un geomorfologo è abituato a delineare rispetto al risultato che scaturisce dal metodo "pixel based". Si possono evidenziare vantaggi nel sia rispetto alla cartografazione manuale che al riconoscimento automatico di tipo "pixel based". Per quanto riguarda la cartografazione manuale, anche un non esperto geomorfologo può essere coinvolto nell'utilizzo di tale tecnica. Difatti, anche un occhio non esperto ed allenato alla cartografazione di forme ed alla redazione di carte tematiche di tipo geomorfologico, con un minimo di esperienza è di solito in grado di riconoscere la validità di una carta già realizzata da altri. Spesso eliminando dagli elaborati la legenda una carta tematica non è più interpretabile. Si pensi ad esempio alle carte geologiche, dove i colori sono associati ad età e litologia ricavati da ore di rilievo di campagna, analisi di laboratorio ecc.. Nel caso delle carte geomorfologiche, essendo la topografia sempre disponibile come base, eliminando la legenda (come nel caso della cluster

analisi) dai gruppi cartografati, è ancora possibile, sovrapponendo le curve di livello, interpretare il significato di ciò che è stato cartografato. Un fondovalle, una cresta, un versante di faglia (anche se come detto, termini indicanti l'origine ha no un significato diverso legato al processo e non alla forma), sono ancora riconoscibili anche ad un occhio non troppo esperto. La carta realizzata in automatico serve quindi a differenziare i gruppi e demanda a chi la osserva di capire quali siano queste differenze e definire una legenda in base a criteri di tipo morfografico (versante rettilineo, impluvio, talweg ecc..) o di tipo genetico (versante di scarpata di faglia, terrazzo fluviale ecc..). Il contro evidente di dover realizzare diverse elaborazioni utilizzando diverse combinazioni di cluster risulta mitigato dai seguenti vantaggi:

- 1) Il geomorfologo è sollevato da ore di studio su carte topografiche spesso di grandi dimensioni e spesso afflitte dal problema della eccessiva densità delle isoipse. In tal senso, poiché il metodo si rivela molto efficace dal punto di vista computazionale, risulta possibile effettuare molte realizzazioni in breve tempo fino a giungere ad o più elaborati finali fra i quali individuare quelli che evidenziano meglio le forme di interesse.
- 2) Il metodo è caratterizzato da una maggiore oggettività rispetto ai metodi di analisi tradizionale. Può essere di ausilio anche ad uno studioso di geomorfologia non esperto, come ad esempio uno studente, perché in genere la fase di valutazione del risultato è più semplice della fase di individuazione e cartografazione. In pratica equivale a leggere una carta geomorfologica senza legenda.
- 3) Una volta estratte le forme di interesse, è possibile passare subito alla fase di analisi specifica (ad esempio di statistica spaziale) poiché il dato è già in formato digitale ed all'interno dell'ambiente GIS.

Per quanto riguarda l'efficacia rispetto al metodo “pixel based” si è visto come in effetti risulta superiore nel riuscire a cartografare le forme piuttosto che i



singoli elementi del versante. Le aree interpretabili come forme sono più omogenee, di maggior estensione ed in generale si nota un pattern (inteso questa volta non come variabile multivariata ma come distribuzione spaziale delle classi e dei pixel) molto meno “caotico”. Tale risultato è ottenuto anche con una buona velocità di esecuzione poiché la complessità computazionale del K-means a parità di estensione dell'area di studio, è proporzionale al numero di variabili utilizzate (e la versione fuzzy è ancora più lenta a causa del calcolo della matrice dei valori membro) ed essendo di fatto necessari solo due variabili nella prima fase di clustering e tre nella seconda, la complessità computazionale non risulta eccessiva.

### 7.3 Cartografazione supervisionata

Nella classificazione supervisionata basata sulla massima verosimiglianza, spesso si utilizza l'assunto che le classi siano governate da una distribuzione di tipo gaussiano. Ciò rende necessario l'utilizzo di variabili la cui distribuzione segua tale modello. Dalla distribuzione normale univariata sostituendo alla varianza la matrice di covarianza ed i vettori delle variabili e delle rispettive medie si ottiene la distribuzione normale multivariata

$$p(x|\mu, \Sigma) = \frac{1}{(2\pi)^{N/2} |\Sigma|^{1/2}} e^{-\frac{1}{2}(\mathbf{x}-\mu)^T \Sigma^{-1}(\mathbf{x}-\mu)}$$

Dove  $\mathbf{x}$  e  $\mu$  sono rispettivamente i vettori delle osservazioni della variabile casuale ed il vettore della media,  $\Sigma$  è la matrice di covarianza delle osservazioni,  $e$  è il numero di Nepero ed  $N$  è la dimensione del vettore, per cui si ottiene una matrice di covarianza  $N \times N$ .

Per classi caratterizzate da matrici di covarianza diagonali, la distanza statistica tra una classe ed un vettore  $\mathbf{x}_i$  è data dalla distanza di Mahalanobis:

$$Dm^2 = (\mathbf{x}_i - \mu_i)^T \Sigma^{-1} (\mathbf{x}_i - \mu_i)$$

Dove la matrice di covarianza  $\Sigma = E[(\mathbf{x} - \boldsymbol{\mu})(\mathbf{x} - \boldsymbol{\mu})^T]$  dove  $E$  indica la media.

Per calcolare la matrice di covarianza è necessario calcolare le medie e varianze/covarianze che sono estratte dai vettori delle variabili appartenenti ad un data set definito come “training dataset”. Si utilizza a tale scopo un certo numero di raster, ognuno rappresentate un livello informativo, da cui verranno estratti i vettori degli attributi. Il “training dataset” è definito tramite l'utilizzo di maschere realizzate tramite delineazione manuale delle aree considerate appartenenti a specifici elementi del paesaggio, da cui verranno estratti i vettori per il calcolo della matrice di covarianza. La delineazione delle aree di interesse è stata effettuata tramite il comando GRASS “r.digit” che permette di realizzare delle maschere raster. Queste aree sono da considerare, come spesso indicato in inglese nella terminologia utilizzata nel campo del remote sensing “ground truth”, ovvero aree di cui si è sicuri l'appartenenza ad una determinata classe (nel nostro caso una particolare forma del paesaggio). Naturalmente un conto è il considerare affidabile il riconoscimento manuale di una determinata area, un altro l'efficacia dei vettori delle variabili scelte utilizzate per il calcolo della distanza statistica. Infatti nel caso della classificazione supervisionata (come d'altronde anche per il clustering), la fase di selezione delle caratteristiche, vista il gran numero di parametri estraibili da un DTM e quindi le numerose combinazioni possibili, imporrebbe di utilizzare un approccio di tipo “tryal and error”. Difatti una semplice analisi del coefficiente di correlazione tra le variabili permette di verificare, tramite l'uso di tabelle, se le variabili sono incorrelate o correlate a due a due. Un basso coefficiente di correlazione per tutte le variabili calcolato in questo modo non permette di verificare se le variabili siano “statisticamente indipendenti” (Webb 2002, pp. 432-444). La distribuzione congiunta può non seguire una distribuzione gaussiana anche se le variabili prese ad una ad una sono caratterizzate da una distribuzione normale.

A tal fine è stato realizzato l'algoritmo “r.mvdist” che calcola la distanza di

Mahalanobis. L'algoritmo richiede l'utilizzo di una maschera. Tale maschera può essere prodotta tramite il comando di GRASS `r.digit`, o importata altrimenti. Questa maschera serve ad individuare i pixel che fanno parte del “training dataset”. L'algoritmo realizzato, sfruttando questa maschera, calcola la media, la matrice di covarianza, l'inversione della stessa e quindi calcola la distanza Mahalanobis per tutte le celle del raster analizzato. Essendo la distanza di Mahalanobis caratterizzata da non avere un valore limite superiore, risulta uno strumento utile all'identificazione di elementi che possono essere mappati con diversi gradi di incertezza. La distanza di Mahalanobis è infatti una distanza statistica, ed equivale alla distanza dalla media espressa in numero di deviazioni standard del caso univariato. Tramite distanza statistica è quindi possibile definire l'incertezza della appartenenza alla classe definita dal “training dataset”.

Una volta calcolato il raster della distanza, si può utilizzare la distanza interna massima (cioè la distanza massima di un vettore appartenente al “training dataset” dalla propria media) o utilizzare un approccio più euristico, basato sulla valutazione visiva del risultato ottenuto in base a diversi valori di distanza.

La distanza di Mahalanobis è stata utilizzata come strumento di individuazione delle forme di origine strutturale presenti nel raster A. Nell'area sono presenti versanti di faglia e forme mature quali faccette triangolari e trapezoidali che sono ben caratterizzate e diffuse ma che, per caratteristiche di pendenza e curvatura, possono facilmente essere confuse con forme simili, ma a scala differente, quali i terrazzi fluviali, che sono gli elementi più estesi nell'area in questione. A tale scopo si è rilevata utile un'analisi multiscala di almeno una parte dei parametri morfometrici utilizzati nella classificazione.

Le forme sopra descritte sono state digitalizzate al fine di ottenere una maschera per il “training dataset”. E' stato quindi utilizzato un sottoinsieme casuale delle forme cartografate per calcolare la distanza di mahalanobis.

I parametri morfometrici utilizzati sono stati: pendenza, curvatura del profilo, curvatura tangenziale, curvatura massima e curvatura minima. Uno dei problemi

nell'utilizzare tali parametri è che essendo calcolati a partire dalle derivate di primo e secondo ordine, sono molto sensibili alle variazioni della superficie. I valori delle derivate variano al variare della scala di analisi. Poiché il metodo si basa sulla distanza statistica tra una distribuzione ed un vettore risulta di fondamentale importanza la scala di analisi. Difatti elementi quali i versanti di faglia, che possono essere considerate forme ascrivibili al macrorilievo, possono presentare, a causa dei processi di erosione, se analizzati a scala più fine, elementi che appartengono al microrilievo. Le superfici di interesse in questo caso dovrebbero essere versanti subplanari che possono invece essere classificati ad esempio come incisioni fluviali se la scala di analisi è troppo grande. Risulta inoltre interessante notare come il problema risulti meno rilevante nel caso della distanza tra due distribuzioni che tra una distribuzione ed un vettore. Difatti il singolo vettore risulta da una misurazione dei parametri di un punto della superficie (o una cella del DTM) per cui definisce la geometria di quel singolo punto. La distribuzione è invece più simile ad uno smoothing dell'area definita dai confini del “training dataset”.

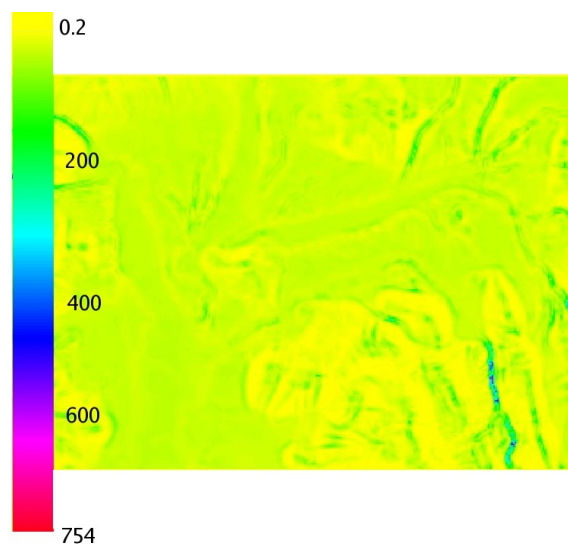
L'utilizzo di una analisi multiscala dei parametri morfometrici, consente di identificare quelle strutture che emergono come elementi del mesorilievo e macrorilievo anche per i vettori. A tal fine i parametri sono stati estratti, tramite il comando GRASS “r.param.scale” utilizzando tre diverse dimensioni della finestra di analisi (3x3, 7x7, 11x11). La scelta in tal caso è di tipo empirico (si suppone che la finestra più grande corrisponda grossomodo alla scala alla quale si manifesta la curvatura dei versanti più estesi; essendo la cella di 40 metri ciò significa un versante di circa 400 metri) e tiene conto del fatto che, oltre che ridurre l'influenza delle frequenze alte, corrispondenti a variazioni a livello del microrilievo, utilizzando più di una finestra di analisi si aggiunge informazione sulle caratteristiche delle celle adiacenti. Il valore della quota di una cella diventa una media pesata di una distribuzione dell'intorno della cella focale. Anche se la forma quadrata della finestra di analisi non segue l'andamento delle forme di

interesse e quindi comprende celle non vi appartengono, la superficie quadratica risulta meno sensibile ai valori estremi rispetto ai filtri medi e di conseguenza anche i valori delle derivate risultano meglio dipendenti dalla reale scala di analisi che dalla distribuzione delle quote all'interno della finestra (Wood 1996). Per calcolare la distanza di Mahalanobis per i versanti di faglia e strutture affini sono stati utilizzati i seguenti parametri morfometrici:

<i>Pendenza 3x3</i>	<i>Pendenza 7x7</i>	<i>Pendenza 11x11</i>
<i>Curvatura del profilo 3x3</i>	<i>Curvatura del profilo 7x7</i>	<i>Curvatura del profilo 11x11</i>
<i>Curvatura tangenziale 3x3</i>	<i>Curvatura tangenziale 7x7</i>	<i>Curvatura tangenziale 11x11</i>
<i>Curvatura massima 3x3</i>	<i>Curvatura massima 7x7</i>	<i>Curvatura massima 11x11</i>
<i>Curvatura minima 3x3</i>	<i>Curvatura minima 7x7</i>	<i>Curvatura minima 11x11</i>

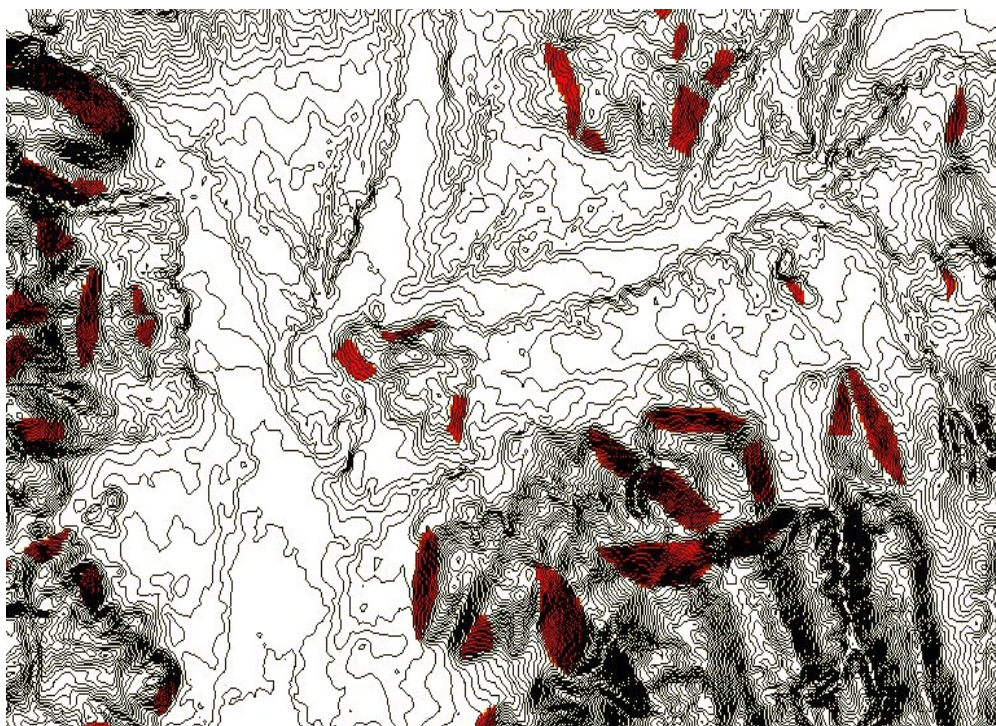
*Tabella 7.6*

La fig. 7.12 mostra gli elementi digitalizzati a partire dalla carta geomorfologica di Brancaccio et alii 2000, mentre in fig 7.13 gli elementi estratti utilizzando una distanza di Mahalanobis pari a 8.

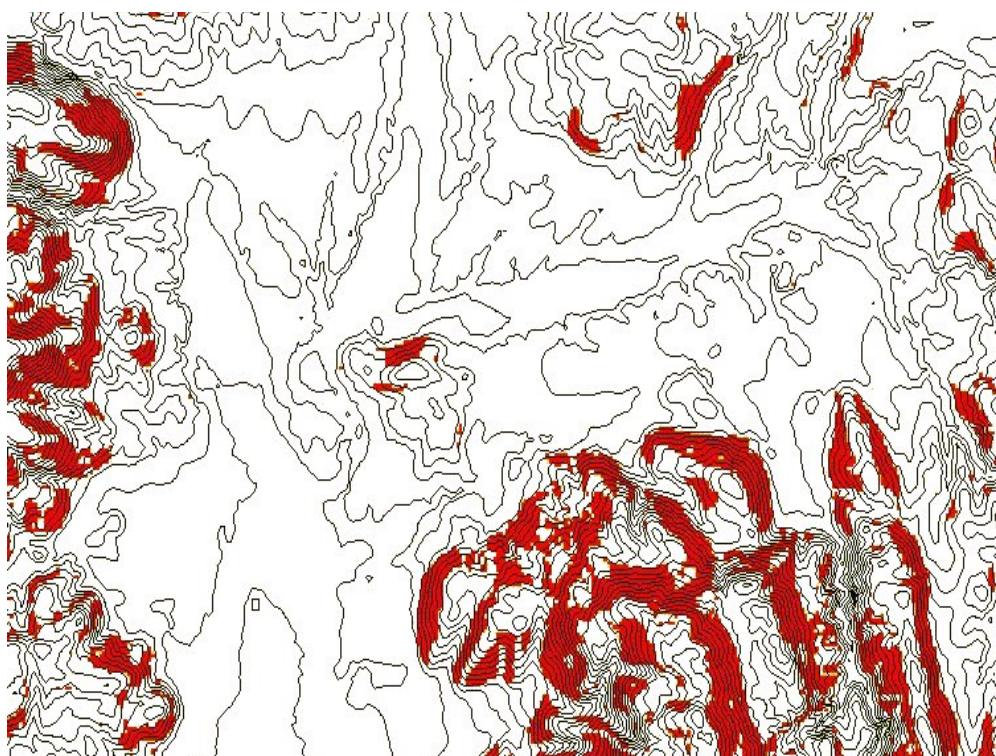


*Fig 7.11 Raster della distanza di Mahalanobis estratto tramite l'algoritmo "r.mvdist" dal raster A (valle del Volturno)*





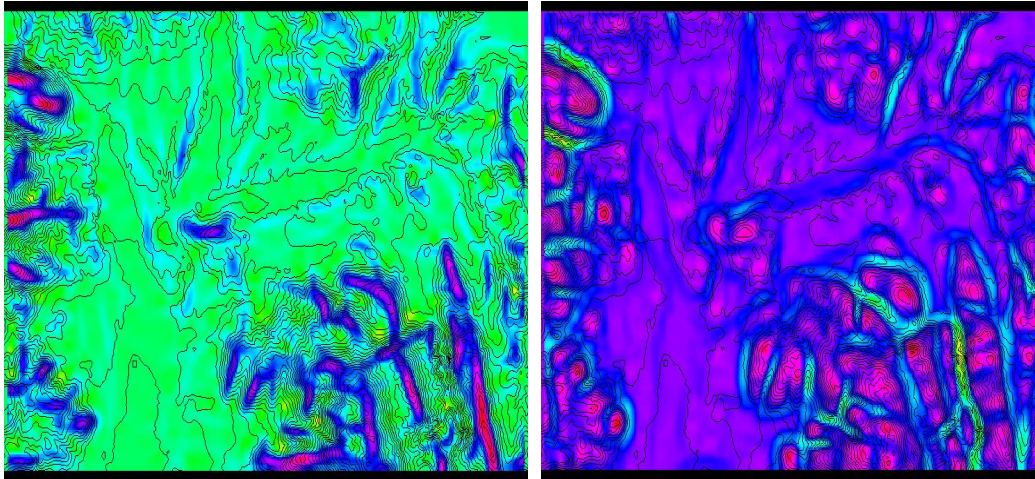
*Fig 7.12 Training dataset digitalizzato dalla carta geomorfologica della valle del Volturno di fig. 7.20*



*Fig 7.13 Versanti di faglia estratti dal raster A utilizzando una distanza di Mahalanobis pari ad 8*



Questo valore equivale a 8 deviazioni standard dalla media e come si vede il risultato rivela una buona corrispondenza tra gli elementi estratti in automatico e quelli presenti nella carta geomorfologica.



*Fig 7.14 Raster della curvatura massima (a destra) e minima (a sinistra) estratti, tramite il comando di GRASS `r.param.scale`, dal raster A. La scala cromatica va dal verde, valori bassi, passando per il blu, valori intermedi, al rosso porpora, valori più elevati.*

## Discussione

Rispetto al criterio di massima verosimiglianza, l'uso della semplice distanza di Mahalanobis non utilizza due o più distribuzioni poiché si è di fronte all'analisi di una sola classe. Pertanto non esiste un criterio, quale quello di minimo rischio, per assegnare un vettore a quella o quell'altra classe. Si determina semplicemente la distanza di un vettore degli attributi, da una classe definita tramite la matrice di covarianza e la media estratti dal “training dataset”. Nel caso fosse possibile utilizzare due distribuzioni, una “del training set” e l'altra dell'elemento da classificare, si potrebbe migliorare il risultato della classificazione utilizzando la distanza tra distribuzioni (di cui esistono vari indici noti in letteratura, es. Webb 2002 pp 425-427). Sfortunatamente non esiste un metodo per definire la distribuzione degli elementi da identificare poiché non si conosce a priori la distribuzione geografica dei pixel appartenenti agli elementi non delineati

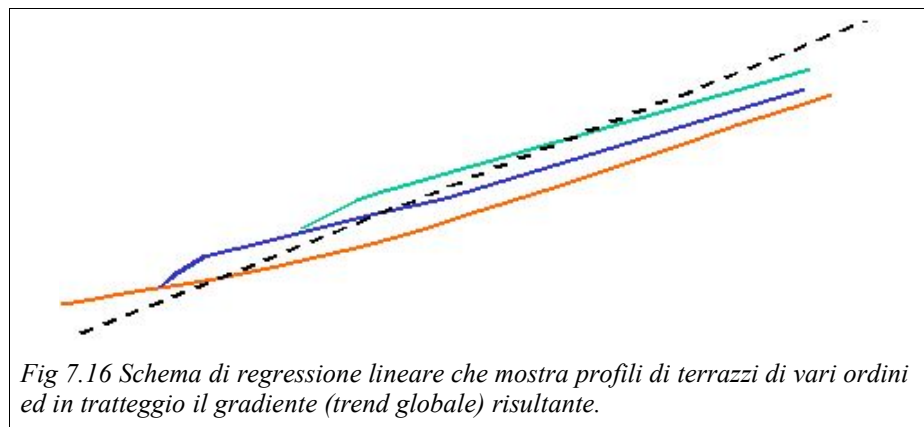
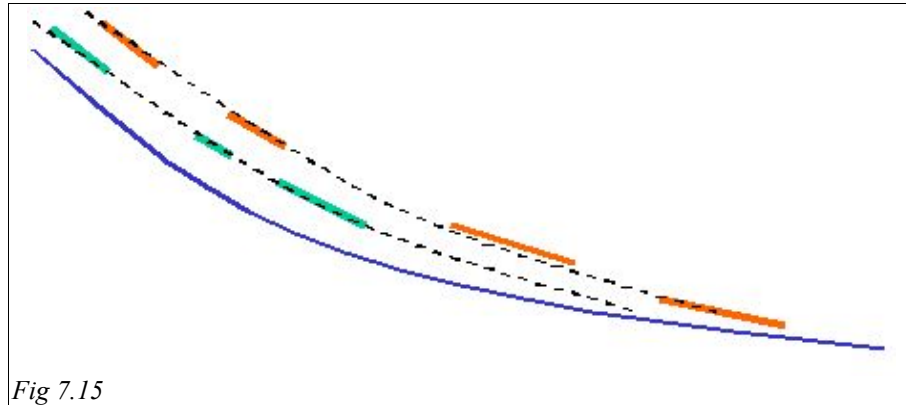
manualmente, cioè proprio ciò che si vuole ottenere da una classificazione supervisionata. Ci si limita allora al calcolo della distanza tra i vettori associati ai singoli pixel del raster ed il “training dataset”. In fig 7.11 si può vedere come la distanza di Mahalanobis varia all'interno dell'area test. Scegliendo un valore limite opportuno è possibile estrarre gli elementi di interesse.

Naturalmente la scelta degli attributi è fondamentale ed in questo caso si è scelto chiaramente in base a quei parametri che dal punto di vista morfometrico meglio caratterizzano le forme in questione. In particolare si può notare in figura 7.14 come varia la distribuzione spaziale della curvatura massima e della curvatura minima estratti con una finestra di analisi 11x11. Essendo i versanti di faglia superfici subplanari, devono essere caratterizzati da bassi valori sia di curvatura minima che di curvatura massima. Ovviamente esistono anche versanti non rettilinei, per questo l'utilizzo degli altri attributi (pendenza, curvatura del profilo ecc..) contribuiscono alla discriminazione di tali forme. L'uso delle finestre a scala multipla, inoltre, vista la presenza di estese forme terrazzate, contribuisce a discriminare tra scarpate riconoscibili come versanti di faglia e scarpate di origine fluviale poiché, benché simili geometricamente, sono elementi del rilievo caratterizzati da scale differenti.



## 7.4 La correlazione ed ordinamento delle superfici

La correlazione dei lembi di terrazzo è un'operazione che può essere facilmente automatizzata se i lembi di uno stesso ordine presentano la stessa correlabilità altimetrica lungo estesi tratti del fondovalle.



E' il caso, ad esempio, della valle del fiume Colorado, dove sono presenti terrazzi di origine prevalentemente erosionale, in cui l'individuazione dei vari livelli, una volta acquisito il DTM dell'area di studio, può essere effettuata semplicemente tramite una riclassificazione delle quote, operazione definita in molti casi come "slicing", ovvero una tecnica impiegata nell'analisi delle

immagini che permette di separare, tramite l'individuazione di un valore soglia, gli elementi di interesse dal cosiddetto sfondo che comprende gli elementi di non interesse. In casi di questo tipo, anche laddove vi fossero livelli non continui spazialmente, risulta relativamente semplice sofisticare la ricerca automatica nel caso il gradiente che caratterizza i vari livelli terrazzati risulti praticamente nullo. Se sussiste tale condizione, è possibile utilizzare la combinazione di quota e pendenza per suddividere i livelli in maniera piuttosto semplice. A tal fine si è utile illustrare un esempio pratico basato sull'utilizzo della map-algebra (cfr. paragrafo 4.5). Una volta estratta la pendenza dal DTM, si può procedere nel modo seguente:

*r.mapcalc raster\_superfici="if(pendenza<pendenza\_max, quota, null())"*

dove “superfici” è il raster risultante, “pendenza” è il nome del raster delle pendenze estratto dal DTM denominato “quota” e assume il valore di “quota” se la pendenza è minore di *pendenza\_max*, o un valore di “no data” nel caso opposto. *Pendenza\_max* è un valore di pendenza massima dei lembi di terrazzi ricavata precedentemente, per esempio mappando un certo numero di terrazzi. Una volta calcolati il valore massimo, media e deviazione standard, tale valore può essere ricavato o dalla pendenza massima di tale campione o un valore ricadente all'interno di *n* deviazioni standard dalla media, in modo da evitare di includere anche scarpate poco acclivi all'interno del processo di classificazione.

Il raster risultante “superfici” sarà allora costituito da lembi isolati che, nel caso il gradiente dei terrazzi sia prossimo a zero, risultano già ordinati poiché i lembi appartenente ad uno stesso ordine sono caratterizzati dalla stessa quota. In pratica i lembi risultano suborizzontali. Tale metodo non è applicabile nel caso i terrazzi presentano un gradiente maggiore (fig 7.15), per cui anche i corrispondenti lembi residui non potranno essere caratterizzati dallo stesso livello altimetrico non essendo suborizzontali ma inclinati secondo la vergenza del

terrazzo.

Quando esiste una vergenza dei terrazzi, un tipo di analisi che viene spesso utilizzata ai fini dell'individuazione dei trend delle superfici, è quella di individuare la superficie di regressione che meglio si adatta all'andamento altimetrico, previa rimozione delle superfici con elevato angolo di pendenza dal DTM. Le aree ad elevata pendenza, come rilievi, fianchi vallivi, forre ecc.. difatti costituiscono dei valori estremi che data la sensibilità del metodo dei minimi quadrati, non permettono di individuare il trend associato ai soli terrazzi. Utilizzando quindi una regressione multipla, basata sulla quota e sulle coordinate, si può tentare di identificare parametri di inclinazione e vergenza delle superfici di interesse.

Il trend individuato è di tipo lineare se si utilizza una regressione lineare, o può essere non lineare se si utilizza un polinomio di grado superiore. Sottraendo i valori del DTM a quelli della superficie di regressione, si ottiene una versione riscalata delle quote i cui ogni valore della cella del raster appartiene a superfici caratterizzate dallo stesso trend, e separate verticalmente. Si possono individuare superfici tramite il loro valore costante e la loro separazione verticale dovuta alla presenza di scarpate fluviali. Spesso però tale metodo fallisce nel separare correttamente gli ordini in sistemi vallivi dove l'energia del rilievo ha dato origine a sistemi terrazzati molto più complessi dal punto di vista plano-altimetrico. Nel caso l'antico fondovalle fosse caratterizzato da una gradiente, le successive fasi di erosione produrranno terrazzi che seguono questo gradiente mentre il nuovo fondovalle, e successivamente gli ordini di terrazzi più recenti, potranno essere caratterizzati da gradienti differenti. E' il caso delle aree di studio analizzate in questa tesi, dove sia la varietà della distribuzione spaziale delle quote del rilievo da cui si è originato il sistema di terrazzi, sia la varietà geolitologica hanno dato origine a sistemi di terrazzi la cui individuazione e correlazione risulta essere molto più complessa. Nel caso infatti di superfici dalla vergenza variabile lungo il bacino fluviale, il metodo della regressione multipla,

che utilizza il metodo dei minimi quadrati, individuando la superficie che meglio si adatta ai punti dati (best fit), per polinomi di basso grado, restituisce un trend dell'insieme delle superfici di vario ordine (fig 7.16). Nel caso invece di polinomi molto convoluti, si ottiene l'effetto indesiderato di un adattamento eccessivo alle superfici, passando di fatto da una analisi del trend ad una interpolazione vera e propria che si avvicina molto alla originale distribuzione altimetrica del DTM originale e di fatto inutile al fine di separare i terrazzi. Spesso i vari ordini sono caratterizzati dallo stesso gradiente, ma risulta comunque impossibile suddividerli tramite una semplice riclassificazione delle quote, a causa della estensione delle superfici. La regressione multipla individua sempre comunque il trend globale dell'insieme delle superfici, non potendo conoscere a priori la separazione verticale delle stesse. In ogni caso i terrazzi fluviali sono superfici caratterizzate da ampie sovrapposizioni del range altimetrico tra i vari ordini. Nelle analisi seguenti, al fine dell'individuazione delle superfici si utilizza un criterio puramente geometrico ma basato su di un criterio differente da quello delle superfici di regressione. In tale contesto le varie informazioni ausiliare, quali la stratigrafia, la correzione delle superfici duplicate causa presenza di faglie, la cui localizzazione potrebbe fornire indicazioni utili sull'individuazione dei limiti geografici entro cui tali superfici sono state dislocate, ed altri dati ancillari utili alla individuazione dell'età relativa dei vari ordini di livelli, non vengono utilizzate poiché lo scopo è quello di sperimentare un metodo che consenta di verificare se sia possibile effettuare la correlazione delle superfici in automatico senza ulteriori livelli informati. Il vantaggio principale dell'uso di algoritmi del tipo sviluppato risulta evidente se confrontato con i metodi tradizionali che sfruttano l'analisi dei profili topografici soprattutto dal punto di vista dei tempi di elaborazione.

## 7.5 Algoritmo sviluppato

Lo scopo iniziale per cui è stato sviluppato l'algoritmo "r.terraces" è quello di consentire in automatico di produrre mappe che accorpino (correlandoli) i lembi residui di un medesimo ordine di terrazzi fluviali e che distingua, se esistono, terrazzi di ordini diversi. Esso è stato in origine ideato per l'utilizzo in aree dove le forme terrazzate sono predominanti rispetto ad altre tipologie di superfici a basso angolo di pendenza, quali –ad esempio- le superfici di spianamento le paleosuperfici di tipo appenninico (solo a tratti subpianeggianti), che possono essere caratterizzate da una dislocazione verticale, escursioni altimetriche e cambiamenti delle direzioni di pendenza molto più accentuati che non i terrazzi fluviali. Tuttavia tale metodo si è rilevato utile anche per analizzare, previa opportune fasi di pre-processamento del DTM, la distribuzione degli ordini di generazioni di superfici ad angolo di pendenza più elevato e più variamente dislocate. Superfici di questo ultimo tipo possono essere ad esempio i glacis vallivi sospesi che si ritrovano a varie quote e la cui distribuzione sia altimetrica che di pendenza è molto variabile nell'ambito del sistema bacino idrografico in cui si rinvencono. I GVS sono caratterizzati da pendenze mediamente più elevate ed interpretabili, data la loro distribuzione lungo entrambi i fianchi di una valle, come relitti di antichi fianchi vallivi.

L'algoritmo è stato perciò testato ed ottimizzato per funzionare in aree dove le superfici correlabili più caratteristiche sono dei terrazzi fluviali, generalmente caratterizzati da separazioni verticali tra un ordine ed il successivo anche di soli pochi metri, le cui superfici di separazione, ossia le scarpate fluviali, sono elementi caratterizzati da forte rottura di pendenza e spesso buona continuità spaziale lungo i margini del terrazzo.

Il tentativo di automatizzazione sviluppato tende sostanzialmente ad emulare quello che è un procedimento tipico dell'approccio umano tradizionale alla problematica : lavorando su carte topografiche, individuato un lembo di terrazzo

fluviale in un certo tratto di una vallata, esplorare la medesima vallata verso monte e verso valle alla ricerca di altre forme analoghe raccordabili. Ossia cercare altri lembi di paesaggio suborizzontali (o inclinati circa come quello di partenza) ed altimetricamente raccordabili tra loro e con quello di partenza.

L'algoritmo sviluppato a tal fine, "r.terraces", utilizza un metodo iterativo basato sul concetto di crescita zonale (fig 7.17) che, partendo da una cella specifica del DTM, effettua una scansione dell'intero raster e ridefinisce le quote in maniera tale da evidenziare lembi di superfici che siano geometricamente correlabili tra di loro (accordanti nel senso di cui in premessa).

Un assunto di base del metodo è che, lavorando su paesaggi policicliche che conservino tracce di più ordini di terrazzi, esista una separazione altimetrica non ambigua tra i vari ordini. Non ambigua qui significa che essa è tale da poter essere discriminata (distinta) dai dislivelli che invece sono insiti nei terrazzi. La distinzione è basata ovviamente sulla pendenza più che sui puri e semplici dislivelli; nel senso che la crescita zonale di un terrazzo (ossia il riconoscimento centrifugo da parte del programma di nuove nuove aree su cui estendere il riconoscimento del terrazzo stesso) si arresta laddove la scansione comincia a trovare pendenze che stanno fuori dal range assegnato alla superficie di cui sta cercando le prosecuzioni (incrementi di quota per unità di passo orizzontale (distanza) superiori ad una soglia prefissata)

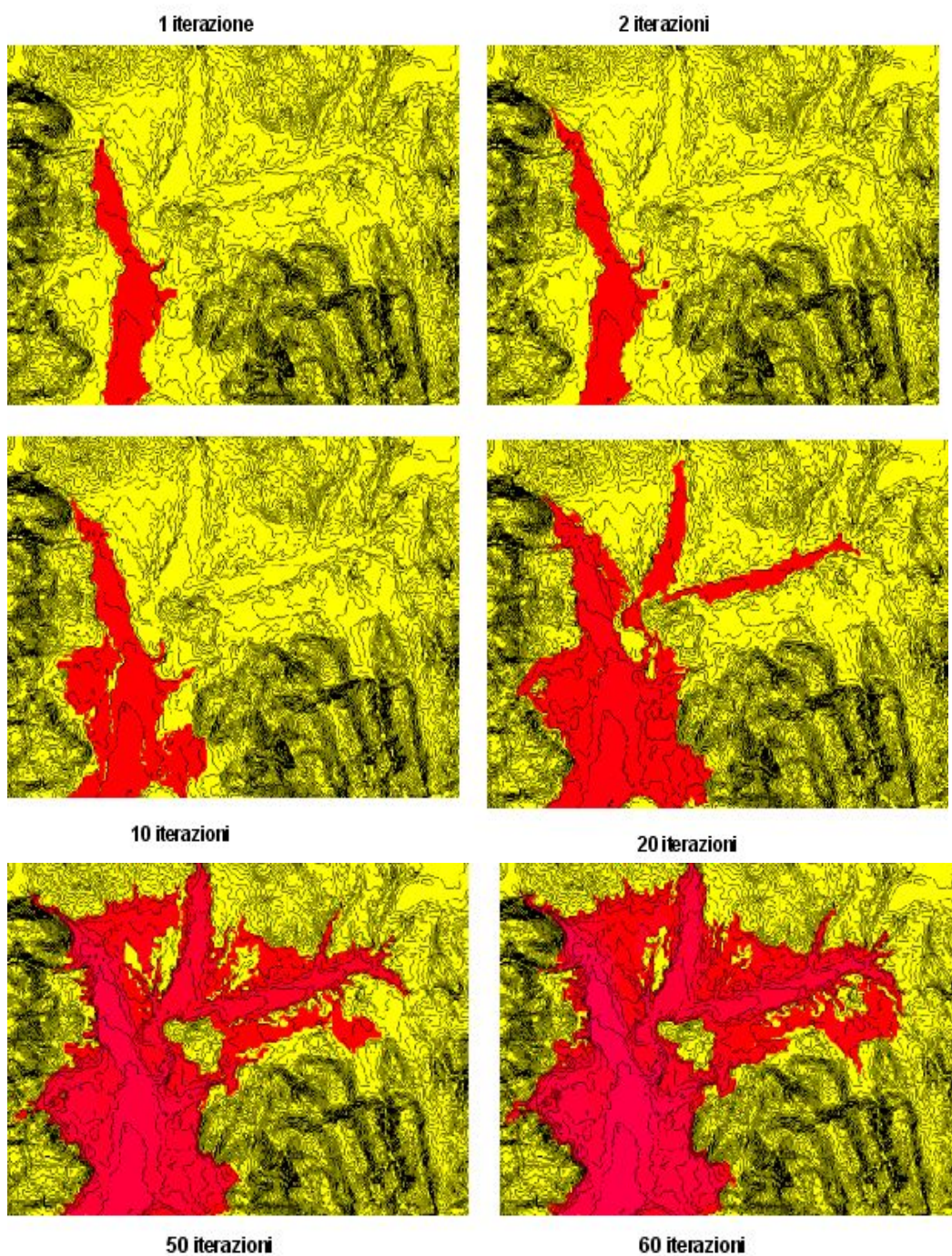
Nel caso di paesaggi in cui le scarpate tra i vari ordini di ripiani abbiano acclività molto modesta (confondibile con pendenze entro i lembi e magari dovute –come spesso accade- a rimodellamenti tardivi) l'algoritmo produrrà risultati poco soddisfacenti.

Il prodotto consiste in un nuovo raster che in base al livello di tolleranza altimetrica imposto dall'utente, ed alla geometria del sistema vallivo e dei singoli lembi di terrazzo, consente di ottenere un valore numerico, qui definito "indice gerarchico" che permette di correlare lembi anche distanti.

Tale indice può essere utilizzato per suddividere i lembi tramite una tecnica

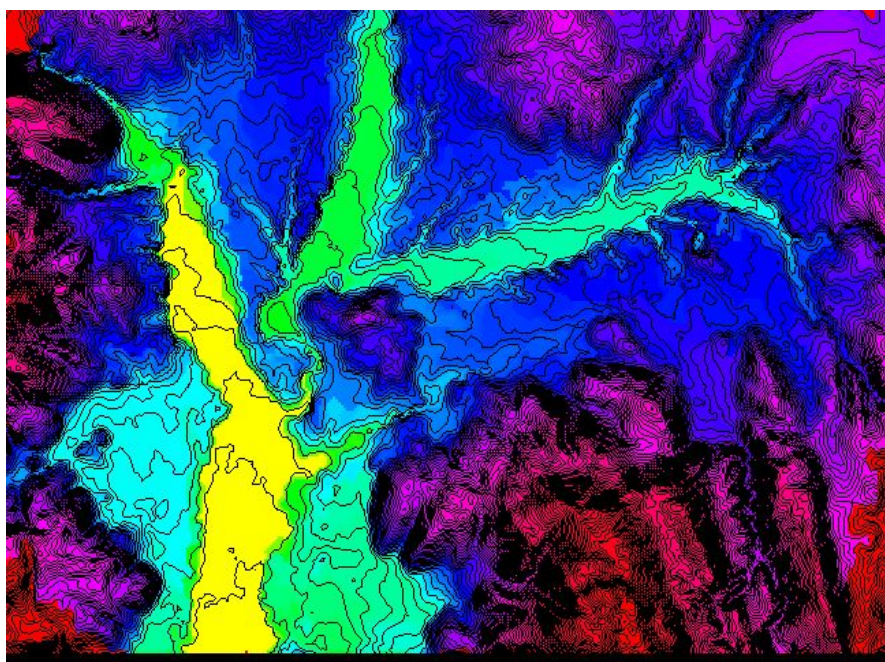


di partizionamento automatico quale il clustering iterativo (es. k-means o fuzzy c-means)



*Fig 7.17*





*Fig 7.18 Indice gerarchico generato per il raster A tramite l'algoritmo r.terraces*

L'algoritmo “r.terraces” implementa una prima fase di scansione della superficie individuando le aree che, al di sotto di un certo angolo di pendenza, definito da una variabile di tolleranza che viene utilizzata per individuare le scarpate associate ai margini interni ed esterni dei lembi di terrazzo. In pratica le celle adiacenti sono processate, a partire da un seme iniziale identificato tramite le coordinate fornite dall'utente, utilizzando un sistema di crescita zonale. Queste coordinate possono essere ricavate mediante un sistema “punta e clicca” implementato nello script Tcl/Tk del sistema a finestre di GRASS, mentre il codice dell'algoritmo è in linguaggio C. Il sistema di selezione tramite cursore consente, visualizzando qualsiasi raster dell'area di studio - ad esempio il raster delle pendenze, di scegliere invece che il punto altimetricamente più basso, un'altro punto conveniente come ad esempio una cella appartenente all'attuale fondovalle. Questa soluzione si rileva migliore rispetto a quella di utilizzare come seme la cella del raster altimetricamente più bassa in assoluto poiché, nel caso si utilizzasse un DTM che comprende vari bacini idrografici, questa potrebbe



ricadere in un bacino al di fuori dell'area di interesse. Inoltre non ha importanza che la cella seme sia realmente la più bassa altimetricamente poiché il criterio della tolleranza altimetrica fa sì che sia sufficiente utilizzare una qualsiasi cella in un'area a bassa pendenza, ad esempio nel fondovalle attuale del bacino che si vuole analizzare. L'algoritmo realizzato funziona secondo il seguente schema:

*Fase di crescita zonale:*

Una volta identificate le coordinate di partenza e quindi la cella iniziale come descritto precedentemente

- 1) Inizializza il contatore delle iterazioni,  $k=0$ , e l'indice gerarchico,  $i=0$ , e classifica la cella iniziale come “processata”. Si inizializza il valore del numero di celle processate  $N(k)=1$ .
- 2) Per ogni cella dell'intorno della cella iniziale si calcola la differenza altimetrica. Al di sotto del valore di tolleranza altimetrica, le celle del raster vengono considerate appartenenti allo stesso lembo di superficie ed identificate come celle processate. In pratica quando il dislivello tra una cella adiacente è inferiore al valore di tolleranza tale cella viene aggregata alle celle già processate ed identificata anch'essa come cella processata e si incrementa la variabile che identifica il numero di celle processate  $N(k)=N(k)+1$ . La cella non viene aggregata ed identificata come processata nel caso il dislivello superi il valore di tolleranza.
- 3) Si continua analizzando tutte le celle classificate come già processate con i criteri definiti al punto 2
- 4) La crescita, cioè il processo di aggregazione, si arresta quando, per tutte le celle identificate come processate fino a quel momento, non è possibile individuare ulteriori celle adiacenti il cui dislivello sia inferiore al valore di tolleranza.
- 5) Si incrementa la quota di tutte le celle identificate come processate di un valore pari alla tolleranza altimetrica.
- 6) Si incrementa il valore del contatore  $k=k+1$

7) Si ripete la fase 2 con i nuovi dati altimetrici. In questo modo, essendo stata incrementata la quota, lungo il bordo della zona le differenze altimetriche tra pixel adiacenti saranno inferiori e alcune differenze tra celle potranno risultare al di sotto del valore limite. Il codice iterativo consente quindi di aggregare nuove celle precedentemente non aggregate perché il dislivello risultante era maggiore del valore di tolleranza imposto. Se con questi nuovi valori altimetrici, durante la fase 3 non risultano aggregate nuove celle, cioè  $N(k-1) = N(k)$  si ripete la fase 5. Altrimenti per ogni nuova cella aggregata si utilizza un valore di indice gerarchico  $i=i+1$

*Fase di postprocessamento:*

Una volta che tutti i pixel dell'area di studio sono stati processati, l'algoritmo prevede una fase di postprocessamento che serve a filtrare le aree anomale. Infatti è possibile, a causa del diverso rapporto plano-altimetrico tra le celle che, durante la crescita zonale, una cella A adiacente ad una cella B, alla fine una indice gerarchico superiore rispetto alla cella A anche se la quota della cella B è inferiore. Ad esempio è facile che scarpate fluviali, lungo il margine esterno, siano identificate con un indice superiore rispetto alla superficie sovrastante per terrazzi fluviali divergenti rispetto al fondovalle. Il codice quindi, al fine di eliminare gli errori di classificazione, scansiona tutti i pixel in cerca di queste anomalie e le corregge secondo il criterio che un pixel posto a quota superiore non può avere un ordine gerarchico "i" inferiore rispetto a superfici adiacenti poste a quote inferiori. Tale procedura decrementa di uno l'ordine - a partire da quello più elevato - per celle adiacenti che non rispettino tale criterio, ed in maniera iterativa fintanto che sussistono celle adiacenti anomale.

## **7.6 Correlazione ed ordinamento dei lembi di terrazzi fluviali**

Il raster A è stato preprocessato tramite il comando `r.fill.dir` di GRASS, al fine di eliminare le depressioni isolate (cfr. paragrafo 3.6) presenti nel DTM interpolato. Tale preprocessing risulta indispensabile poiché l'algoritmo `r.terraces` si basa sull'analisi dei dislivelli tra celle adiacenti. Le depressioni isolate producono una contropendenza che l'algoritmo non è in grado di gestire. In tal caso la crescita zonale si blocca in maniera anomala, soprattutto se le depressioni sono situate in corrispondenza di valli strette (in pratica la larghezza della depressione è paragonabile alla larghezza della valle). Dopo il preprocessing l'algoritmo riesce a procedere per tutte le celle del raster senza bloccarsi. L'area test rappresentata dal raster A è stata quindi stata processata dall'algoritmo `r.terraces` utilizzando un valore di tolleranza pari a 0,8 metri. La suddivisione in ordini dei lembi di terrazzo è stata prodotta utilizzando un metodo automatico. Innanzitutto le aree ad elevata pendenza sono state scorporate in modo da isolare solo le superfici di interesse. Successivamente il raster risultante è stato partizionato in 4 ordini, tramite `k-means` utilizzando l'algoritmo `r.clustering`. Il numero di corretto di cluster è stato individuato dopo vari tentativi per confronto con la carta geomorfologica. La scelta è stata dettata quindi dalle seguenti considerazioni.

I terrazzi di II e III ordine, a causa della bassa risoluzione del DTM non sono sempre correttamente associati all'ordine giusto e comunque sono scarsamente presenti nell'area di studio. L'unico lembo significativo risulta essere il fondovalle del fiume Carpino che però, dal risultato del clustering, risulta inglobato nei terrazzi di II ordine, presumibilmente a causa di una leggera dislocazione tettonica. Dal risultato del clustering inoltre si vede come in effetti esiste la duplicazione dell primo ordine. I terrazzi, in giallo nella figura 7.19, che circondano l'abitato di Isernia, sono dislocati tettonicamente. Altri lembi inglobati nello stesso cluster di questi terrazzi, sono in effetti lembi di superfici sommitali

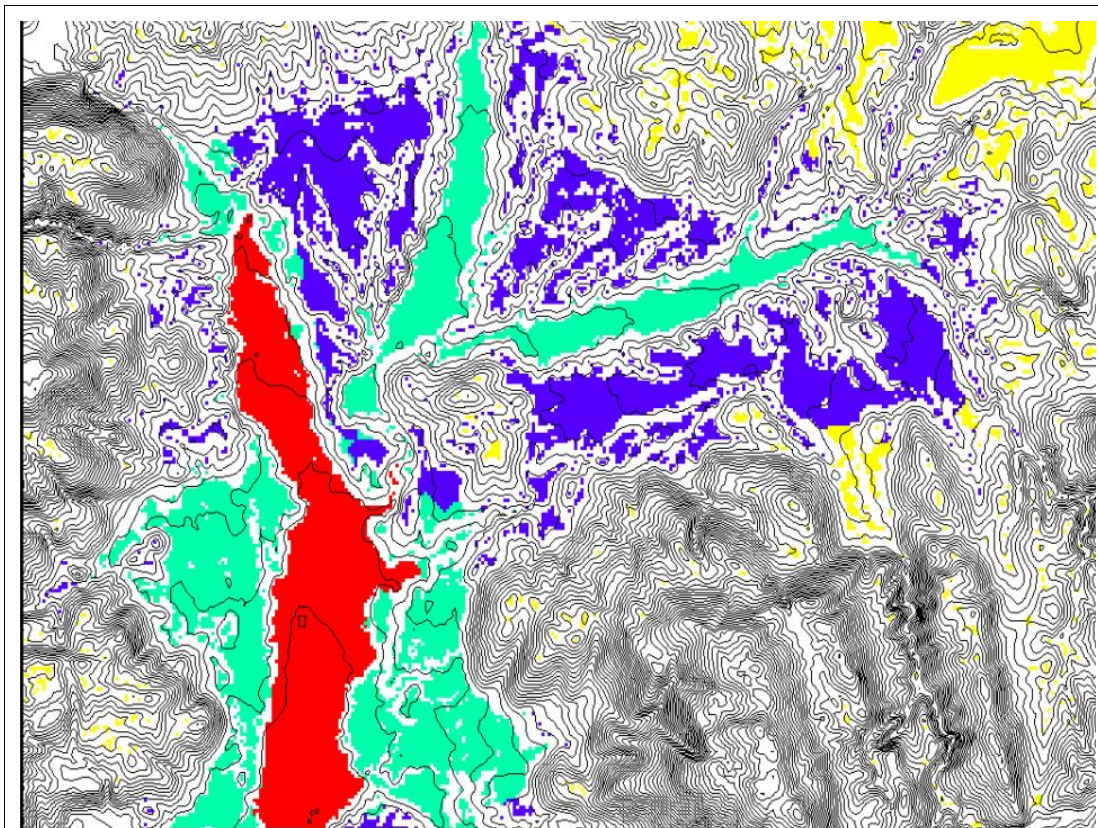


Fig 7.19 Suddivisione tramite *k-means* del raster A in 4 classi di terrazzi fluviali utilizzando l'indice gerarchico prodotto da "r.terraces" e la pendenza. Rosso: fondovalle + terrazzi di IV ordine, Verde: III e II ordine, Blu: I ordine. Giallo: Terrazzi del I ordine dislocati e lembi di supefici relitte.

di spianamento come si può notare dal confronto con la carta geomorfologica di fig 7.20.





## 7.7 Analisi e ordinamento dei glacis

Per individuare l'ordine dei glacis si è utilizzato lo stesso algoritmo impiegato per i terrazzi fluviali. Nel caso dei glacis, essendo caratterizzati da pendenze medie elevate (fig 7.21), si è utilizzato una tolleranza più elevata rispetto a quella utilizzata per separare i terrazzi fluviali. Sono stati provati diversi livelli di tolleranza, poiché la differenza di pendenza tra un glacis e l'altro e anche la differenza di pendenza all'interno dello stesso lembo possono variare abbastanza. Una tolleranza troppo bassa o una tolleranza troppo elevata hanno entrambi lo stesso effetto negativo di non consentire una suddivisione delle quote in base all'ordine di generazione. Nel primo caso, se la tolleranza è inferiore alla pendenza equivalente delle superfici di glacis, queste verranno suddivise in maniera scorretta, producendo diverse fasce altimetriche che includono lo stesso ordine di glacis. Nel secondo caso, essendo la tolleranza superiore alla pendenza delle superfici glacis e pari o superiore alle superfici di raccordo, queste verranno inglobate in un'unica fascia altimetrica e quindi non separate in ordini. Purtroppo il problema non è di facile soluzione poiché in pratica lembi anche dello stesso ordine possono presentare pendenze differenti. Ciò implica necessariamente di trovare un compromesso tra valore di tolleranza e superfici correttamente suddivise in ordini. A tal fine è stato analizzato l'andamento planimetrico delle superfici mappate manualmente e l'andamento del teorico margine esterno risultante dall'applicazione dell'algoritmo.

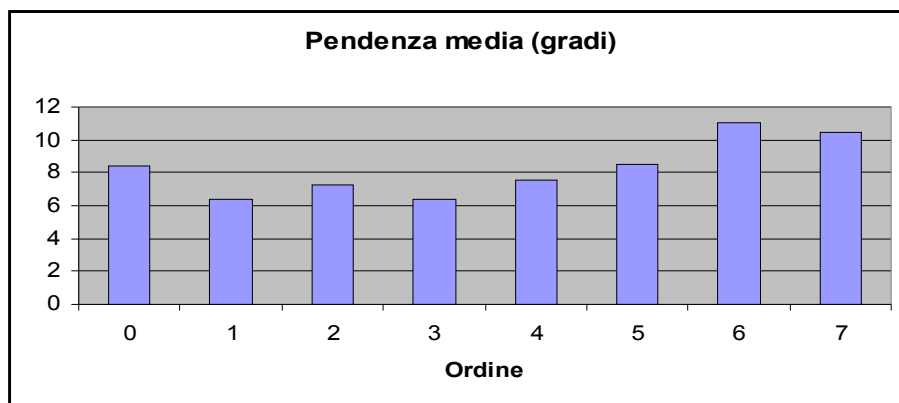


Fig 7.21 Pendenza media per ordine di glacis 1-7 ordini di GVS, 0=glacis di incerta sede

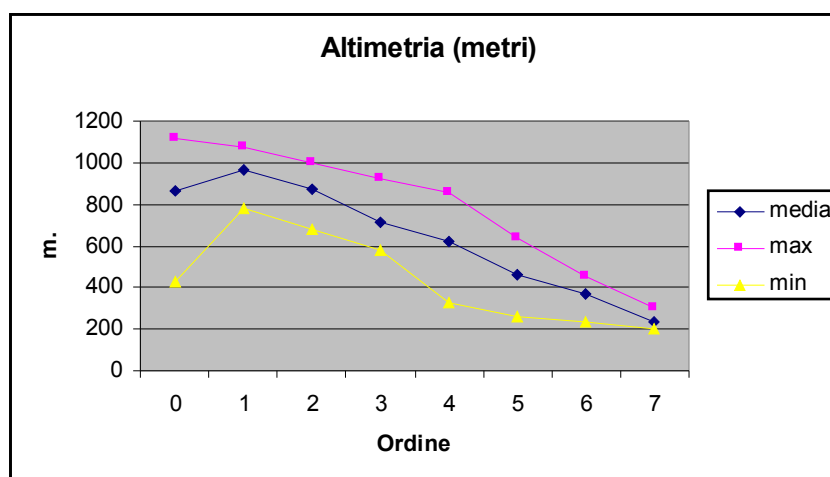


Fig 7.22 Quota media minima e massima calcolata per i lembi di figura 6.3

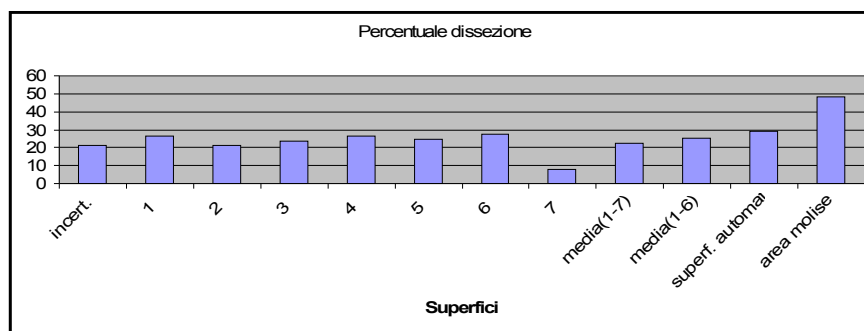


Fig 7.23 Percentuale di superficie dissecata per ogni ordine di glacis e la media per il territorio della regione Molise

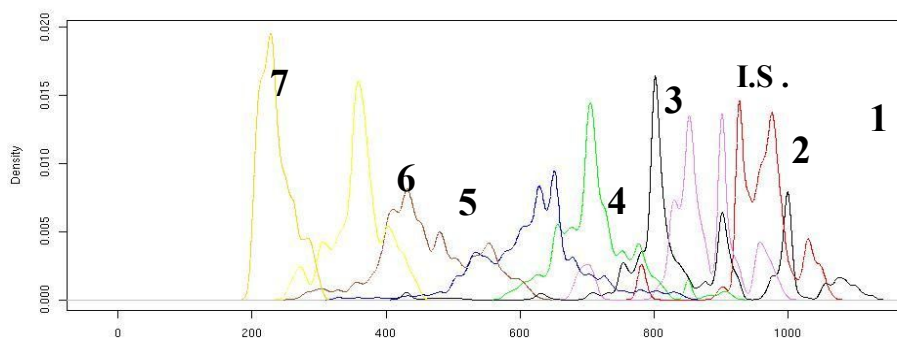


Fig 7.24 Distribuzione altimetrica per ordine di glacis; 1-7 ordini di GVS (I.S.=Glacis Incerta sede)

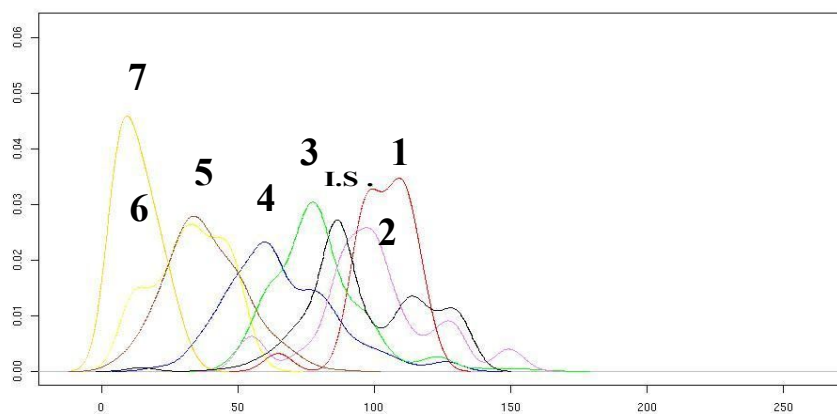


Fig 7.25 Distribuzione dell'indice gerarchico ottenuto da "r.terraces" per ordine di glacis; 1-7 ordini di GVS (I.S.=Glacis Incerta sede)



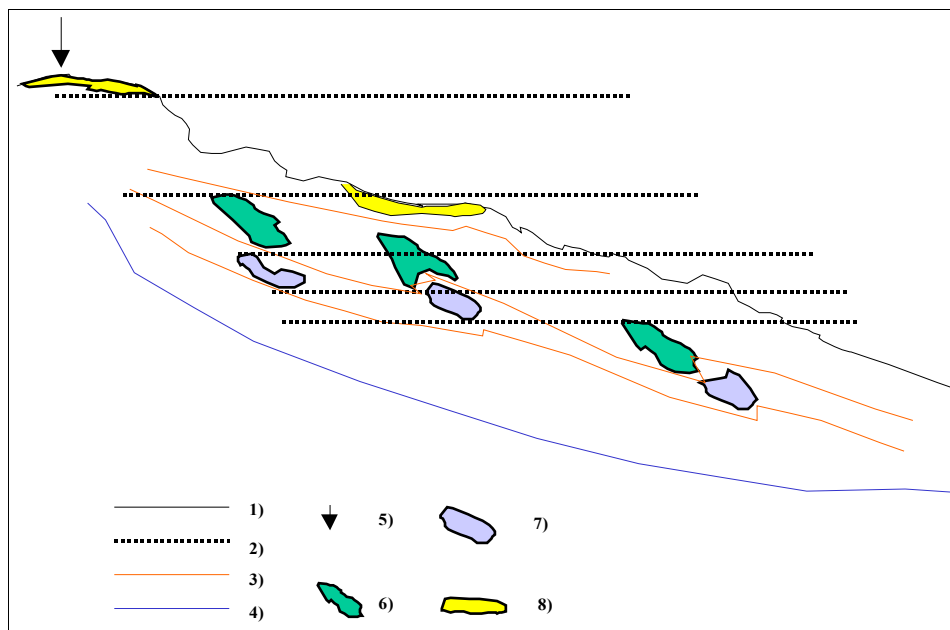
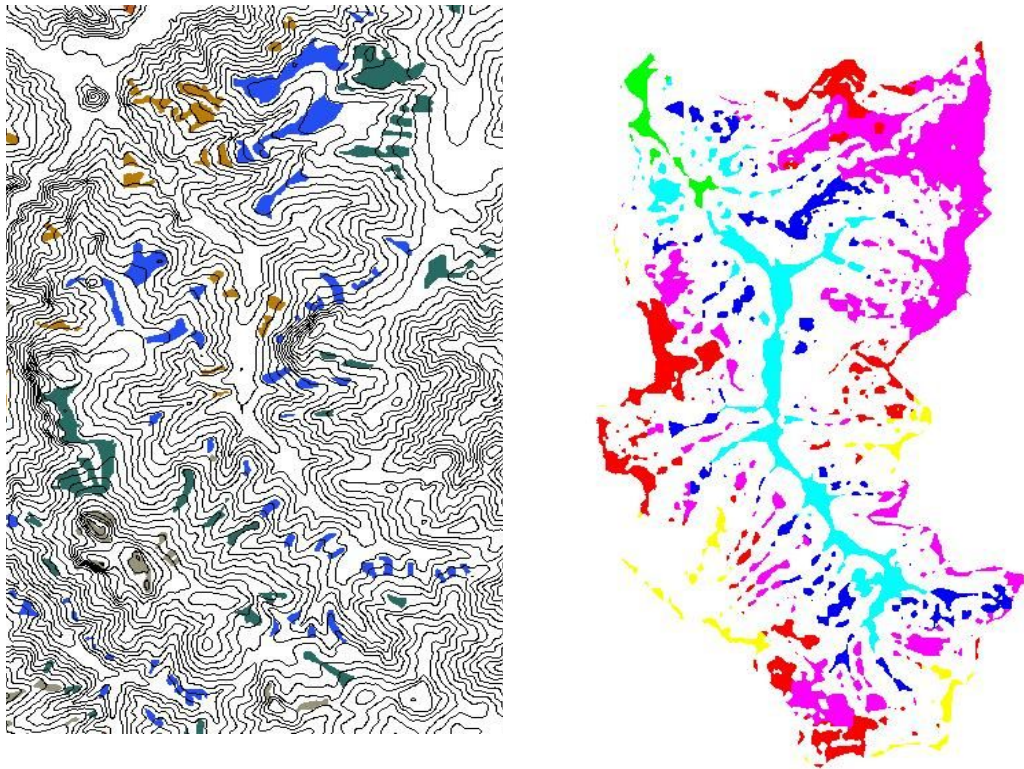


Fig 7.26 Schema teorico della geometria associata ai GVS. 1) Profilo longitudinale del bacino fluviale 2) Fasce altimetriche: evidenziano la sovrapposizione altimetrica di due glacis di differente ordine di generazione 3) Andamento altimetrico dei margini interni ed esterni delle superfici di glacis di vario ordine 4) Fondovalle attuale 5) Spartiacque 6) Glacis più antico 7) Glacis più recente 8) Paleosuperfici sommitali

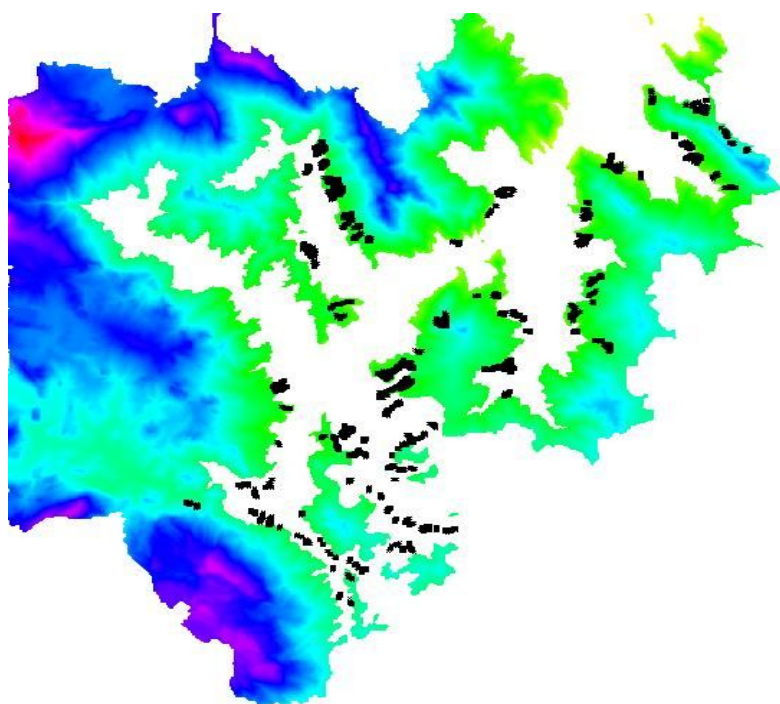
Individuare le superfici di glacis è comunque un problema di non facile soluzione, Ricordando che i glacis dello stesso ordine si rinvencono smembrati ed a distanze notevoli tra loro, è possibile individuare le superfici poste a quota più elevata per ogni singola superficie, seguendo localmente l'andamento delle curve di livello. Se si vuole invece individuare l'intera area che sovrasta tutto l'ordine è evidente che, data la distanza tra un glacis e l'altro, non è possibile utilizzare le quote massime e minime dei glacis dello stesso ordine ciò poiché i glacis si sviluppano altimetricamente seguendo il gradiente della paleovalle. Le fig 7.22 ed 7.24 evidenziano come, per i glacis mappati manualmente, esistono notevoli sovrapposizioni nei valori di quota. In pratica un certo numero di lembi di glacis di un determinato ordine, si rinvencono a quote uguali od inferiori rispetto a glacis di ordine superiore (cioè più giovani). Tramite l'algoritmo "r.terraces" si è ottenuto un nuovo *raster* applicando un livello di tolleranza pari a 2.5 metri. Su questo nuovo *raster* delle quote riscalate (ovvero dell'indice gerarchico ricavato dall'algoritmo) sono state ricalcolate le distribuzioni altimetriche degli stessi

glacis. Il risultato, illustrato in fig. 7.25, mostra come anche in tal caso si notano ampie sovrapposizioni. Nel caso di fig. 7.25 la sovrapposizione riguarda la distribuzione dell'ordine gerarchico prodotto da r.terraces. Ciò significa che purtroppo non è possibile, a partire da tale strato informativo, ottenere una separazione degli ordini mediante un partizionamento automatico (ad esempio tramite k-means). Si è tentato comunque di realizzare, come nel caso dei terrazzi fluviali, una suddivisione in ordini che che rispecchiasse il risultato ottenuto dalla cartografazione manuale (fig 7.27) tuttavia con risultati questa volta non soddisfacenti. L'algoritmo r.terraces non riesce a seguire tutti i margini esterni ed interni dei singoli lembi a causa della estrema variabilità in pendenza degli stessi. Ciò determina il problema della sovrapposizione delle distribuzioni anche di tale indice gerarchico. Non esiste di conseguenza un livello di tolleranza ottimale in grado di fornire una buona separabilità tra gli ordini. La pendenza dei vari lembi difatti varia sensibilmente non solo tra i vari lembi ma anche all'interno dello stesso lembo e ciò implica che l'algoritmo impiegato può solo in maniera approssimativa individuare i margini esterni ed interni associati ai vari ordini. A tal proposito si può però notare che dal punto di vista planimetrico, una volta bucato il DTM originale, l'andamento del presunto margine esterno di un determinato ordine, segue con buona approssimazione l'andamento dei glacis mappati manualmente. L'operazione che consente di bucare il DTM consiste nel selezionare solo le celle del raster che siano superiori ad un determinato valore di indice gerarchico. Per tentativi è possibile bucare, ossia in termini tecnici, assegnare valori di “no data” al DTM fino ad individuare tale margine sovrapponendo al DTM i poligoni dei glacis mappati manualmente. Questo sarà il bordo che divide le aree di “no data” dalle aree ricoperte da valori altimetrici. In fig 7.28a e 7.28b il DTM è stato bucato utilizzando il risultato di “r.terraces” ottenuto specificando due diversi livelli di tolleranza (2.5 e 3 metri). Sono stati quindi sovrapposti i glacis di ordine 5, rappresentati da poligoni, verificando che il margine fosse coerente con l'andamento degli stessi. Tale margine esterno

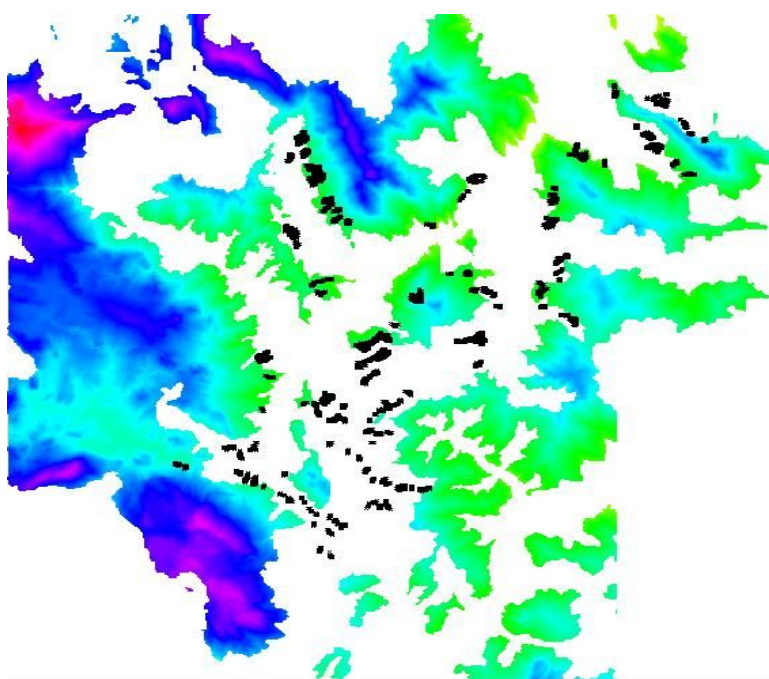
risulta pur sempre approssimato ma comunque coerente con l'andamento planimetrico dei lembi (anche se non per tutti i lembi cartografati manualmente).



*Fig. 7.27 A sinistra glacies mappati manualmente derivati da (Aucelli 1999) del bacino del fiume Vella (raster C). A destra tentativo di ordinamento automatico delle stesse superfici in 4 ordini tramite clustering dell'indice gerarchico prodotto da "r.terraces" (il fondovalle risulta spezzato in due ordini)*



A



B

*Fig 7.28 Due casi del raster D bucato utilizzando l'indice gerarchico generato dall'algoritmo r.terraces. In fig B si è utilizzata una tolleranza di 2.5 metri. In fig A si è utilizzata una tolleranza di 3 metri. I poligoni sovrimposti corrispondono ai glacis di ordine 5. L'andamento planimetrico risulta coerente con il margine ma non per tutti i lembi.*

## ***C.IV Conclusioni e sviluppi futuri***

Il lavoro mostra come sia possibile introdurre sensibili miglioramenti ai metodi che, partendo dalla sola informazione estraibile dai modelli digitali del terreno, puntano ad automatizzare almeno alcuni passaggi del riconoscimento e conseguente mappatura di forme del paesaggio secondo criteri geomorfologici.

Se gli elementi che si vogliono analizzare sono individuati manualmente, il DTM (e gli strumenti GIS utilizzati per le analisi) risulta strumento di immediata utilità potendo impiegare tutte le tecniche di analisi quantitativa note in letteratura agli elementi di interesse (geomorfometria specifica). Caso più complesso è quello dalla cartografazione automatica. I DTM possono essere impiegati per estrarre automaticamente attributi morfometrici lungo la superficie e quindi elementi geometrici di base del rilievo quali elementi concavi, convessi o angoli di pendenza. Ma per il riconoscimento e la cartografazione di forme composite, quali le forme descritte in A.II, ed in paesaggi policiclici e/o poligenici, per differenziate in ordini di generazione, i DTM devono essere elaborati tramite algoritmi che seguono criteri analoghi a quelli impiegati dal geomorfologo nell'analisi del rilievo.

Il metodo proposto per la classificazione non supervisionata mostra come anche con l'uso di una tecnica di *clustering* iterativo appositamente modificata, sia possibile realizzare rapidamente una classificazione e quindi una cartografazione di aree che evidenzia forme composite piuttosto che elementi geometrici semplici. Tali elementi sono caratterizzati dagli attributi morfometrici estratti dalle singole celle del DEM che non consentono di realizzare classi separabili nello spazio degli attributi, associate ad forme composite. L'estrazione attributi differenti, in questo caso le medie di alcuni parametri morfometrici, consente di ottenere classi separabili associabili alle forme più complesse. Il risultato finale risulta molto differente da una semplice “cluster analysis”. Difatti risulta molto più simile ad una tradizionale carta geomorfologica.

Benché il metodo nel suo complesso comporti una certa soggettività nel passaggio in cui l'operatore sceglie il numero di *cluster*, la rapidità con la quale è possibile ottenere gli elaborati cartografici rispetto ai metodi manuali, consente di sopperire a tale limite attraverso la sperimentazione di diverse combinazioni di parametri dell'algoritmo (numero di classi e riclassificazione dei layer). I risultati metodologici sinora ottenuti sono molto importanti perché prefigurano la messa a punto di procedure di analisi geomorfologica automatizzata di estrema utilità laddove si tratta di analizzare aree molto ampie. L'individuazione rapida degli indicatori geomorfologici di antichi livelli di base dell'erosione risulta di grande utilità soprattutto laddove solo lo studio della loro distribuzione altimetrica a scala regionale (ad esempio: in più bacini idrografici contigui) può far emergere i trend di ampio raggio legati ai sollevamenti ed inarcamenti tettonici recenti rispetto al "rumore di fondo" dovuto a differenze di erodibilità ed altri fattori locali.

Il fatto che il metodo sviluppato ha positivamente superato i test di applicazione ad aree di grande frammentarietà e complessità geomorfologica, costituisce una garanzia circa la sua piena applicabilità in aree meno articolate.

Per completezza si è implementato nel GIS un algoritmo per il calcolo della distanza di Mahalanobis. Tale algoritmo è stato utilizzato per individuare in maniera supervisionata le forme a controllo strutturale presenti nel raster A (valle del Volturno). Rispetto alla classificazione basata sulla massima verosimiglianza, tale metodo permette di concentrare la ricerca su di un'unica classe ovvero un unico elemento del paesaggio. Infatti la massima verosimiglianza ed altri metodi di classificazione supervisionata, anche se ottimali dal punto di vista della minimizzazione del rischio di errore di classificazione, sono designati per assegnare ogni pixel ad una delle classi presenti nell'area. Il metodo utilizzato invece permette di concentrarsi e quindi mappare rapidamente, singoli elementi del paesaggio. Anche in questo caso è necessario un minimo di esperienza nel valutare il risultato della cartografazione, che è determinata dalla distanza

statistica imposta dall'utente, oltre che dai parametri morfometrici scelti per la classificazione (che però possono essere facilmente dedotti dalle caratteristiche geometriche delle forme di interesse, come nel caso presentato dei versanti strutturali).

Per quanto riguarda la correlazione dei lembi di superfici relitte e l'ordinamento, è stato realizzato un algoritmo che si è dimostrato efficace nell'individuare i vari ordini di terrazzi fluviali in un'area test abbastanza complessa. L'algoritmo individua un indice gerarchico associandolo alle celle del DTM. Tale indice si rivela utile per correlare i lembi e quindi raggrupparli in ordini di generazioni. Il raggruppamento può essere eseguito ad esempio attraverso varie prove di riclassificazione manuale mediante “slicing” dell'indice gerarchico, questo nel caso l'operatore sia in grado di giudicare visivamente l'estensione e la veridicità della separazione ottenuta. E' stato tuttavia proposto e sperimentato il raggruppamento automatico dei lembi tramite *clustering*. Questo metodo necessita da parte dell'operatore solamente di indicare il numero di ordini che si presume siano presenti, e si è dimostrato soddisfacente nel raggruppare i lembi di terrazzi fluviali nell'area test.

Nel caso dei glacis purtroppo il metodo non è risultato particolarmente efficace. Il problema deriva dall'estrema variabilità in pendenza. Ogni lembo è caratterizzato da pendenze differenti non correlabili con l'ordine di generazione. Anche all'interno dello stesso lembo la pendenza risulta molto variabile. Poiché l'algoritmo *r.terraces* utilizza un valore di tolleranza fisso, risulta più adeguato all'individuazione degli ordini di superfici caratterizzate da pendenze più omogenee come i terrazzi fluviali.

Nel complesso possiamo affermare che, malgrado diversi limiti dovuti sia all'utilizzo di algoritmi che essendo di tipo iterativo possono portare a differenti soluzioni ed ai problemi insiti nella estrazione delle derivate della superficie legati alla qualità dei DTM (cfr. 3.7), i metodi proposti risultano comunque più oggettivi rispetto ai metodi tradizionali di cartografazione. Esistono metodi

ancora più oggettivi per la classificazione automatizzata del paesaggio (cfr. Dikau 1995) ma che sono soggetti a rigide e prestabilite scelte dei parametri oltre che numerosi altri problemi dovuti ad esempio allo schema di analisi a finestre mobili ed ovviamente soggetti anch'essi alla stessa problematica dell'estrazione dei parametri morfometrici dal DTM.

### **Sviluppi futuri**

Affinare l'algoritmo r.terraces: a tale scopo si richiede l'integrazione di un metodo che consenta di modificare in maniera adattiva il valore di tolleranza altimetrica in base alla variabilità del gradiente delle superfici processate. Questo migliorerebbe sensibilmente la correlazione tra terrazzi fluviali su più lunga distanza. E' stata osservata (cfr. 7.7) una buona coerenza dal punto di vista planimetrico, ma non altimetrico, tra l'indice gerarchico e l'andamento di un numero significativo di lembi di glacis ricavati da carta geomorfologica. Ciò fa sperare che con ulteriori miglioramenti risulti possibile ottenere una maggiore correlabilità altimetrica anche per tali morfologie più complesse.

Migliorare la tecnica di classificazione non supervisionata: questo può essere realizzato tramite l'utilizzo di tecniche che analizzino il risultato finale ottenuto; ad esempio i bordi delle superfici tramite analisi di Fourier, per caratterizzare meglio le forme anche dal punto di vista planimetrico.



## Bibliografia

Ahnert F. (1984) *Local Relief and the height limits of mountain ranges*, American Journal of Science, vol. 284

Amato A. (1995) *Le paleosuperfici dell'Appennino campano-lucano in rapporto all'evoluzione tettonica neogenico-quadernaria*, Tesi di Dottorato in Geologia del Sedimentario, Università degli Studi di Napoli Federico II

Aucelli P.P.C., Amato A., Cinque A. (2003) *The long-term denudation rate in the South Appenines Chain (Italy): a GIS-aided estimation of the rock volumes eroded since middle, Pleistocene time*, Quaternary International, 01-102, 3-11

Aucelli P.P.C., Cinque A., Esposito A. (2005) *L'uso di tecniche di classificazione non supervisionata per la mappatura automatizzata di forme a partire da modelli digitali del terreno*, Geogr. Fis. Dinam. Quat., Suppl. VII, 35-40

Aucelli P.P.C., Cinque A., Roskopf C. (2001) *Geomorphological map of the Trigno basin (Italy): explanatory notes*, Geogr. Fis. Dinam. Quat. 24, 3-12.

Aucelli P.P.C. (1999), *Analisi morfodinamica del bacino del fiume Trigno: distribuzione entità e cause dei fenomeni di alveo e di versante*, Tesi di Dott. di Ricerca in Scienze Ambientali: uomo e ambiente. XII Ciclo - Università degli Studi del Molise.

ArcView 3.2 (1993) Environmental Systems Research Institute inc.

ArcView (1996) Using the Arcview Spatial Analyst (Manual), ESRI

Ball G.H. & Hall D.J. (1967) *A clustering technique for summarizing multivariate data*, Behavioral Science, 12, 153-155

Beatson R.K., Cherrie J.B., Mouat C.T. (1999) *Fast fitting of radial basis functions: Methods based on preconditioned GMRES iteration*, Advances in Computational Mathematics, 11, 253-270

Birot P., Lhenaff R., Monteux P. (1958) *Recherches sur la limite septentrionale des glaciers d'érosion en roches tendres*, Memories et documents - Centre de Recherches et Documentation Cartographiques et Géographiques, vol. 6, 35-57.

Bivand R. S. (2000) *Using the R statistical data analysis language on GRASS 5.0 GIS database files*, Computers & Geosciences, 26, pp. 1043-1052.

Bivand R. S., Gebhardt A. (1998) *Implementing functions for spatial statistical analysis using the R language*, 38° Congress of the European Regional Science Association, Vienna, 28 August - 1 September 1998.

Bivand R. S., Neteler M. (2000) *Open Source geocomputation: using the R data analysis language integrated with GRASS GIS and PostgreSQL data base systems*, Proceedings 5th conference on GeoComputation, 23-25 August 2000, University of Greenwich, U.K.

Bivand, R. S. (1999) *Integrating GRASS 5.0 and R: GIS and modern statistics for data analysis*. In: Proceedings 7th Scandinavian Research Conference on Geographical Information Science, Aalborg, Denmark, 1999, pp. 111-127

- Bjorneberg D. L., Aase J.K., Trout, T.J. (1997) *WEPP model erosion evaluation under furrow irrigation*, Paper no. 972115, ASAE Ann. Int. Meeting, Minneapolis, MI
- Bonin O. & Rousseaux F. (2005) *Digital Terrain Model Computation from Contour Lines: How to Derive Quality Information from Artifact Analysis*, *Geoinformatica* 9:3, 253–268
- Bovic A.C. (1991) *Analisis of multicannel narrow-band filters for image texture segmentation*, *IEEE Transcation on Signal procesing*, 39(9)
- Brancaccio L., Di Crescenzo G., Rosskopf C., Santangelo N., Scarciglia F. (2000) *Carta geologica dei depositi quaternari e carta geomorfologica dell'alta valle del Fiume Volturno (Molise, Italia meridionale) note illustrative*, *Il Quaternario* 13(1/2), 81-94
- Brown D.G, Lusch D.P., Duda K.A. (1998) *Supervised classification of types of glaciated landscapes using digital elevation data*, *Geomorphology* 21, pp 233-250
- Buja A., Swayne D. F., Littman M., Dean N., Hofmann H.(2001). *XGvis: Interactive Data Visualization with Multidimensional Scaling*, *Journal of Computational and Graphical Statistics*
- Burrough P., Gaans (van) P.F.M., MacMillan R.A. (2000) *High-resolution landform classification using fuzzy k-means*, *Fuzzy Sets and Systems* 113 (2000) 37-52
- Callant J.C., Hutchinson M.F. (1997) *Scale dependence in terrain analysis*, *Matematics and Computers in Simulation*, 43, 313-321
- Carr J.C., Beatson R. K., Cherrie J.B., Mitchell T. J., Fright J.B., McCallum J.B., Evans T.R. (2001) *Reconstruction and Representation of 3D Objects with Radial Basis Functions*, *ACM SIGGRAPH 2001*, Los Angeles, CA, pp 67-76, 12-17 August 2001
- Carter J.R. (1988) *Digital Representations of Topographic Surfaces: An Overview*, in *American Congress on Surveying and Mapping and American Society for Photogrammetry and Remote Sensing*, *Technical Papers* 5:54-60
- Carrara A., Cardinali M., Guzzetti F., Reichenbach P. (1995) *GIS-based techniques for mapping landslide hazard*. (<http://deis158.deis.unibo.it>)
- Castiglioni G.B. (1986) *Geomorfologia (seconda edizione)*, UTET Torino
- Cebecauer T., Hofierka J., Suri M. (2002) *Processing digital terrain models by regularized spline with tension: tuning interpolation parameters for different input datasets*, *Proceedings of the Open source GIS - GRASS users conference 2002*
- Chorowitz J., Parrot J., Taud H, Hakdaoui M. , Rudant J.P. , Rouis T. (1995) *Automated pattern-recognition of geomorphic features from DEMs and satellite images*, *Zeitsch. für Geomorph.*, Supplementband 101
- Comincioli V. (1991) *FORTRAN 77: introduzione e applicazioni numeriche*, McGraw-Hill
- Comincioli V. (1995) *Analisi Numerica: metodi, modelli, applicazioni*, McGraw-Hill
- Conrad O. (2002) DiGeM – Software for Digital Analysis  
<http://www.geogr.uni-goettingen.de/pg/saga/digem/home.html>

- Coulthard T.J., M.J. Kirkby & M.G. Macklin (1997) *Modelling hydraulic, sediment transport and slope processes, at a catchment scale, using a cellular automaton approach*, conference GeoComputation 1997 & SIRC 1997
- Dainelli G. (1926) *Discordanza tra orografia e idrografia nell'Appennino campano*, Riv. Geogr. It., 33, 75-81
- Davis W.M. (1899) *The Geographical cycle*, Geogr. Journal, 14(A), 481-503
- Dikau R. (1989) *The application of a digital relief model to landform analysis in geomorphology*. in Three dimensional applications in GIS, Raper, J. ed., Taylor and Francis, London, pp.51-77
- Dikau R., Brabb E.E., Mark R.M. (1995) *Morphometric landform analysis of New Mexico*, Zeitsch. für Geomorph. Suppl. Bd. 101, 109-126
- Evans, I. S. (1972). *General geomorphometry, derivatives of altitude, and descriptive statistics*. in Chorley, R. J. ed. Spatial Analysis in Geomorphology, Methuen, London. pp.17-90.
- Evans, I.S. (1980) *An integrated system of terrain analysis and slope mapping*, Zeitsch. für Geomorph. Suppl. Bd. 36, 274-295
- Eyton, R. J. (2005) *Unusual display of DEMs*, Cartographic Perspectives, 50, 7-23
- FarField Technology Ltd. (2003), *FastRBF toolbox manual*
- Ferri M., Fainelli A. (1996) *L'interferometria SAR per la creazione di modelli digitali del terreno*, SIFET, 4, 131-147
- Fielding E.J. (2000) *Morphotectonic evolution of the Himalayas and Tibetan Plateau*, in Geomorphology and Global Tectonics, John Wiley & Sons
- Florinsky I.V.(1996) *Quantitative topographic method of fault morphology recognition* Geomorphology 16 , 103-119
- Forsyth D. & Ponce J. (2003) *Computer Vision - A Modern Approach*, Prentice-Hall
- Foster G.R. & Wischmeier W.H. (1974), *Evaluating irregular slopes for soilloss prediction*, Transactions of the ASAE, 12, 305-309
- Friedman J.H. (1987) *Exploratory projection pursuit*, Journal of the American Statistical Association, 82\_249-266
- Goodchild, M. F., Mark, D. M. (1987). *The fractal nature of geographic phenomena*, Ann. Assoc. Of Am. Geogr. 77(2), pp.265-278.
- GRASS (2000) *Geographic Resources Analysis Support System*. <http://grass.itc.it>
- Guth P.L. (1999) *Quantifying and visualizing terrain fabric from digital elevation models*, international conference on Geocomputation 1999
- Guzzetti F. & Reichenbach P. (1994) *Toward The definition of topographic divisions for Italy*, Geomorphology, 11, 1994, p. 57-75
- Hammond E.H. (1964a) *Analysis of proprieties in land form geography: An application to broad-scale land form mapping*, Ann. Assoc. Of Am. Geogr. 54:11-19

- Haralick, R.M. (1983) *Ridge and valley on digital images*, Computer Vision, Graphics and Image Processing, 22(1), pp.28-38
- Harrison J.M., Lo C.,(1996) *Pc-based two-dimensional discrete fourier transform programs for terrain analysis*, Computer&Geosciences vol 22 n. 4 ,419-424
- Horn B.K.P. (1981) *Hill shading and the reflectance map*, Proceedings of the IEEE, Vol. 69-1:14 ,47
- Horton R.E. (1945) *Erosional development of streams and their drainage basins; Hydrographysical approach to quantitative morphology*, Geol. Soc. America Bull.,56:275-370
- Hovius N. (2000) *Macroscale process system of mountain belt erosion*, in Geomorphology and Global Tectonics, John Wiley & Sons
- Hutchinson M.F. (1988), *Calculation of hydrologically sound digital elevation models*, Proceedins of the Third International Symposium on Data Handling, Sydney. International Geographical Union, Columbus, 177-133
- Irvin B.J., Ventura S.J., Slater B.K. (1997) *Fuzzy and isodata classification of landform elements from digital terrain data in Pleseant Valley Wisconsin*. Geoderma, 77(2-4)
- Isaaks H.E, Srivastava M.R. (1989) *An introduction to applied geostatistics* , Oxford University Press
- Jain A.K., M.N. Murty, P.J. Flynn (1999) *Data clustering: a review*, ACM computing Surveys, Vol. 31 No. 3
- Jenson, S.K., Domingue J.O. (1988). *Extracting topographic structure from digital elevation data for geographic information system analysis*, Photogrammetric Engineering and Remote Sensing 54 (11), 1593–1600.
- Jones R. (2002) *Algorithms for using a DEM for mapping catchment areas of stream sediment samples*, Computers &Geosciences 28, 1051 – 1060
- Jordan G. (2002) *Morphometric Analysis and tectonic interpretation of digital terrain data: a case study*, Earth Surface and Landforms, 28, 807-822
- Kidner D., Dorey M., Smith D. (1999) *What's the point? Interpolation and extrapolation with a regular grid DEM*, international conference on Geocomputation 1999
- King C.A.M. (1976) *Landforms and Geomorphology. Concepts and History*, Benchmark paper in Geology, 28, Hutchinson & Sons
- Kirkpatrick S., Gelatt C.D. jr., Vecchi M.P. (1983) *Optimization by simulated annealing*, Science, 220, 671-680
- Lee T.W. (1998) *Independent Component Analysis*, Kluwer Academic Publisher
- Lehmann O. (1933) *Morphologische Theorie der Verwitterung von Steinshlagwanden*, Viertel. der Naturforsch. Ges. Zurich, 83-236
- Lentini F. (1979) *Le Unità Sicilidi della Valle d'Agri (Appennino lucano)*, Geologica Rom., 18, 215-224.

- Lou W. (2001) *LANDSAP: a coupled surface and subsurface cellular automata model for landform simulation*, Computer & Geosciences 27, 363-367
- Mandelbrot B.B. (1982) *The Fractal Geometry of Nature*, W.H. Freeman, New York
- Marchini J. L. <marchini@stats.ox.ac.uk>, Heaton C. <chrisheaton99@yahoo.com>, & Ripley B. D. <[ripley@stats.ox.ac.uk](mailto:ripley@stats.ox.ac.uk)>
- Martz, L.W., Garbrecht J. (1998) *The treatment of flat areas and depressions in automated drainage analysis of raster digital elevation models*, Hydrological Processes 12, 843-855
- Mayer L. (2000) *Application of digital elevation models to macroscale tectonic geomorphology*, in Geomorphology and Global Tectonics, John Wiley & Sons
- Merritts D., Vincent K., Wohl E. (1994) *Long river profiles, tectonism, and eustasy: A guide to interpreting fluvial terraces*, Journal of Geophysical Research vol. 99 n. B7
- Mitas L., Mitsova H., Brown W.M., Astley M. (1996) *Interacting fields approach for evolving spatial phenomena: application to erosion simulation for optimized land use*, Proc. of the III. Int. Conf. On Integration of Environmental Modeling and GIS, Santa Fe (NM)
- Mitas, L., Mitsova, H., (1999), *Spatial Interpolation*. In: P.Longley, M.F. Goodchild, D.J. Maguire, D.W.Rhind (Eds.), Geographical Information Systems: Principles, Techniques, Management and Applications, Wiley.
- Mitsova H, Hofierka J. (1993) *Interpolation by regularized spline with tension: II. Application to terrain modeling and surface geometry analysis*, Math. Geol. 25, 657-669
- Mitsova H., Mitas L. (1993), *Interpolation by Regularized Spline with Tension: I. Theory and implementation*, Mathematical Geology, 25, 1993, 641-55.
- Mitsova H., Mitas L., Brown W.M., Warren S. (1997) *Multi-dimensional GIS environment for simulation and analysis of landscape processes*, Paper No. 973034, ASAE Annual International Meeting, Minneapolis, Minnesota.
- Mitsova H., Hofierka J., Zlocha M., Iverson R. L. (1996a) *Modeling topographic potential for erosion and deposition using GIS*, Int. Journal of Geographical Information Science, 10(5), 629-641.
- Moore, I.D., Grayson, R.B., Ladson, A.R. (1991) *Digital terrain modelling: a review of hydrological, geomorphological and biological applications*. Hydrological Processes, 5, 3-30
- Naesset E. (1997) *Estimating timber volume of forest stands using airborne laser scanner data*, Remote Sensing of Environment, 61: 246-253
- Newmann M.W., Sproull F.R. (1979) *Principles of interactive computer graphics*, McGraw Hill
- O'Callaghan J.F. & Mark D.M. (1984). *The extraction of drainage networks from digital elevation data*, Computer Vision, Graphics, and Image Processing 28, 323-344
- O'Sullivan D. & Unwin D.J. (2003) *Geographic Information Analysis*, Wiley

- Ortolani F., Narciso G., Sanzo A. (1975) *Prime considerazioni sulla presenza del Flysch Numidico nell'Appennino sannita*, Boll. Soc. Natur. in Napoli, 84, 31-44
- Outcalt S., Hinkel K., Nelson F. (1994) *Fractal physiography?*, Geomorphology 11, 91-106
- Panizza M. (1992) *Geomorfologia*, Pitagora Editrice Bologna
- Patacca E., Scandone P., Bellatalla M., Perilli N., Santini U. (1992a) *La zona di giunzione tra l'arco appenninico settentrionale e l'arco appenninico meridionale nell'Abruzzo e nel Molise*, Studi Geologici Camerti, vol. spec. (1991/2), CROP 11, 417-441.
- Patacca E., Scandone P., Bellatalla M., Perilli N., Santini U. (1992b) *The Numidian-sand event in the Southern Apennines*, Mem. Sc. Geol. Padova, 43, 297-337.
- Penck W. (1953) *Morphological Analysis of Landforms*, MacMillan, London
- Perlant F. (2000) *Using stereo images for digital Terrain Modeling*, Surveys in Geophysics 21: 201–207
- Peter A. S. , Larisa S. S., Andrew V. M. (2002) *Foundamental quantitative methods of land surface analysis*, Geoderma 107, 1-32
- Pike R.J. (1988) *The geometric signature: Quantifying landslide terrain types from digital elevation models*, Mathematical Geology, 20 (5), pp.491-511
- Pike R.J. (1995) *Geomorphometry – progress, practice and prospects*, Zeithschrift fur Geomorphologie Suppl. Bd. 101, 109-126
- Pike R.J. (2001) *Digital terrain modeling and industrial surface metrology: converging realms*, Professional Geographer, 53(2), 263-274
- Pike R.J. (2002) *A Bibliography of Terrain Modeling (Geomorphometry), the Quantitative Representation of Topography - Supplement 4.0*, U.S.G.S. open file report of-02-465
- Pitas I. (1992) *Digital Image Processing Algorithms*, Prentice Hall
- Quinn P., Beven K., Chevallier P., Planchon O. (1991) *The prediction of hillslope flow paths for distributed hydrological modelling using digital terrain models*, Hydrological Processes 5, 59–79
- Ribolini A. (2000) *Relief distribution, morphology and cenozoic differential uplift in the Argentera Massif (French-Italian Alps)*, Zeitsch. für Geomorph., 44(3), 363-378
- Shapiro, M. & Westervelt J. (1992) *r.mapcalc: An algebra for GIS and Image Processing*, US Army Construction Engineering Research Laboratory, Champaign, IL
- Siami L. L., Braucher R., Bourles D. L. (2000) *Les nucléides cosmogéniques produits in-situ: de nouveaux outils en géomorphologie quantitative*. Bull. Soc. Géol. France, n. 4, 383-396.
- Skidmore A.K. (1989) *A comparison of techniques for calculating gradient and aspect from a gridded digital elevation model*, International Journal of Geographical Information Systems, 3(4), 323-334

- Speight J.C. (1967) *Parametric description of landform*. in Stewart, G. A. (ed), Land Evaluation, Proceedings of the CSIRO/UNESCO symposium, Canberra 26-31, pp.239-250
- Stone R. & Dugundji J. (1965) *A Study of Microrelief - Its Mapping, Classification, and Quantification by Means of a Fourier Analysis*, Engineering Geology, 1(2):89-187
- Strahler A.N. (1968) *Quantitative geomorphology*, in Fairbridge, R.W., ed., The Encyclopedia of Geomorphology, New York, Reinhold, p. 898-912.
- Stralher A.N. (1957) *Quantitative analysis of watershed geomorphology*, Am. Geophys. Union. Trans. 38(6):913-920
- Summerfield M.A. (2000) *Geomorphology and Global Tectonics*, John Wiley & Sons
- Swayne D. F. ,Cook D., Buja A. (1998) *XGobi: Interactive Dynamic Data Visualization in the X Window System*, Journal of Computational and Graphical Statistics 7 (1)
- Talmi A. & Gilat G. (1977) *Method for Smooth Approximation of Data*, Journal of Computational Physics, 23: 93-123
- Tarboton D.G. (1997) *A new method for the determination of flow directions and upslope areas in grid digital elevation models*, Water Resources Research 33 (2), 309–319
- Theodoridis S., Koutroumbas K. (2003) *Pattern recognition 2<sup>nd</sup> edition*, Elsevier
- Thomas R. & Nicholas A.P., (2002) *Simulation of braided river flow using a new cellular routing scheme*, Geomorphology ,43,179-195
- Tippet M.J, Hovius N. (2000) *Geodynamic processes in the Souther Alps, New Zeland*, in Geomorphology and Global Tectonics, John Wiley & Sons
- Vaumas E. (1963) *Sur le caracteristiques morphologiques des versants periglaciaires non fonctionnels de solifluction laminaire ("glacis d'erosion en roches tendres")*. Compte Rendus Hebdomadaries des Seances de l'Accademie des Sciences, vol. 256, pp. 3163-3166.
- Vorosmarty C.J., Fekete B.M., Meybeck M., Lammers R.B.,(2000) *Geomorphometric attributes of the global system of rivers at 30-minute spatial resolution*, Journal of Hydrology 237,17– 39
- Walker J.S. (1988) *Fourier Analysis*, Oxford, New York
- Walker J.S. (1997) *Fourier Analysis and Wavelet Analysis*, Notices of the Amer. Bull. Math. Soc., 44:6, 658-670
- Webb A. (2002) *Statistical Pattern Recognition*, John Wiley and Sons
- Weibel R. & Brändli M. (1995) *Adaptive methods for refinements of digital terrain models for geomorphometric applications*, Zeitsch. für Geomorph. Suppl. Bd. 101, 13-30
- Wilson J.P., Gallant J.C. (2000) *Digital terrain analysis*. In Wilson J. P. and Gallant J. C. (eds) Terrain Analysis: Principles and Applications. New York, John Wiley and Sons: 1-27.

- Wise S.M. (1998) *Digital Terrain Models - Traps for the Unwary*, international conference on Geocomputation 1998
- Wood, J.D. (1996), *The geomorphological characterisation of digital elevation models*, Ph.D dissertation, Geography Department, University of Leicester, UK
- Young A. (1972) *Slopes*, Oliver & Boid
- Zahouani H., Vargiolu R., Loubet J.L. (1998) *Fractal models of surface topography and contact mechanics*, Mathl. Comput. Modelling Vol. 28 n.4-8
- Zatelli P., Antonello A. (2002) *New GRASS modules for multiresolution analysis with wavelets*, Proceedings of the Open source GIS - GRASS users conference 2002
- Zevenbergen L.W., Thorne C.R. (1987) *Quantitative analysis of land surface topography*, Earth Surface Processes and Landforms 12: 47-56.



## Appendice

### Script della shell che implementa il metodo di classificazione non supervisionata

```
#!/bin/bash
#
# 1) produrre i raster delle quote (dem) pendenza (slope) e curvature (es con
v.surf.rst
# o r.slope.aspect)
# 2) lanciare il comando con "sh cls.sh [opzioni]"
# es:
# GRASS> sh clusclump.sh clump1=<raster_dem> clump2=<raster_slope>
clump3=<altro> ...
# clump110=<...> layer1=<profile curvature> layer2=<tangential curvature>
layer3=<altro> ..
# layer10=<...> class1=<numero cluster per il clumping> class2=<numero classi
finali>
# nonoltre=<max numero di iterazioni k-means> output=<raster finale>
#
#
# Lo script dovrebbe utilizzare la pendenza e la quota per realizzare la mappa
dei clump
# Successivamente il clustering si dovrebbe basare su pendenza, curvatura del
profilo
# e curvatura tangenziale. Si può usare altrimenti una combinazione di raster
(fino a 10)
# a scelta
# Nota: La pendenza é consigliabile sempre per il clustering finale
#
# Es: per utilizzare anche la quota nel clustering finale usare
# sh clusclump.sh clump1=<dem> clump1=<slope> layer1=<profc> layer2=<tangcurv>
layer3=<dem>
# e' tuttavia possibile usare anche parametri differenti per il clusterig
iniziale
# infatti clump1=...clump10= sono i raster di input per il clumping
# layer1=....layer10=.... i raster input per la classificazione finale
```

```
declare -a LAYER ;
declare -a CLAYER ;
```

```
nc=0
n=0
for i do
  case $i in
```

```
    clump1=*) CLAYER[1]=`echo $i | sed s/clump1=//`
              nc=`expr $nc + 1` ;;
    clump2=*) CLAYER[2]=`echo $i | sed s/clump2=//`
              nc=`expr $nc + 1` ;;
    clump3=*) CLAYER[3]=`echo $i | sed s/clump3=//`
              nc=`expr $nc + 1` ;;
    clump4=*) CLAYER[4]=`echo $i | sed s/clump4=//`
              nc=`expr $nc + 1` ;;
    clump5=*) CLAYER[5]=`echo $i | sed s/clump5=//`
              nc=`expr $nc + 1` ;;
    clump6=*) CLAYER[6]=`echo $i | sed s/clump6=//`
              nc=`expr $nc + 1` ;;
    clump7=*) CLAYER[7]=`echo $i | sed s/clump7=//`
              nc=`expr $nc + 1` ;;
    clump8=*) CLAYER[8]=`echo $i | sed s/clump8=//`
```

```

nc=`expr $nc + 1` ;;
clump9=*) CLAYER[9]=`echo $i | sed s/clump9=//`
nc=`expr $nc + 1` ;;
clump10=*) CLAYER[10]=`echo $i | sed s/clump10=//`
nc=`expr $nc + 1` ;;

layer1=*) LAYER[1]=`echo $i | sed s/layer1=//`
n=`expr $n + 1` ;;
layer2=*) LAYER[2]=`echo $i | sed s/layer2=//`
n=`expr $n + 1` ;;
layer3=*) LAYER[3]=`echo $i | sed s/layer3=//`
n=`expr $n + 1` ;;
layer4=*) LAYER[4]=`echo $i | sed s/layer4=//`
n=`expr $n + 1` ;;
layer5=*) LAYER[5]=`echo $i | sed s/layer5=//`
n=`expr $n + 1` ;;
layer6=*) LAYER[6]=`echo $i | sed s/layer6=//`
n=`expr $n + 1` ;;
layer7=*) LAYER[7]=`echo $i | sed s/layer7=//`
n=`expr $n + 1` ;;
layer8=*) LAYER[8]=`echo $i | sed s/layer8=//`
n=`expr $n + 1` ;;
layer9=*) LAYER[9]=`echo $i | sed s/layer9=//`
n=`expr $n + 1` ;;
layer10=*) LAYER[10]=`echo $i | sed s/layer10=//`
n=`expr $n + 1` ;;

classi1=*) CL1=`echo $i | sed s/classi1=//` ;;
classi2=*) CL2=`echo $i | sed s/classi2=//` ;;
nonoltre=*) MAXIT=`echo $i | sed s/nonoltre=//` ;;

output=*) OUT=`echo $i | sed s/output=//` ;;

*)
echo ""
echo "Opzione non valida: " $i
esac
done

k=2
com=", "
FF="{CLAYER[1]}"
while [ $k -le $nc ] ; do
FF=$FF$com${CLAYER[$k]}
k=`expr $k + 1`
done

r.clustering input=$FF output=clust1.clusclump k=$CL1 maxiter=$MAXIT

r.mapcalc clust1.clusclump="clust1.clusclump+1" > /dev/null

r.clump -q input=clust1.clusclump output=clump1.clusclump

k=1
kk=1
while [ $k -le $n ] ; do
F="caverage.clusclump"$k
C="clump1.clusclump"

```

```

r.fpstats dat=${LAYER[$k]} cove=$C stat=$F tota=tt.clusclump
echo $MESSAGE${LAYER[$k]}
k=`expr $k + 1`

done

r.volume -q data=caverage.clusclump1 clump=clump1.clusclump
site_list=site.clusclump >/dev/null

MESSAGE="Converting: "
k=1
kk=1
while [ $k -le $n ] ; do
F="caverage.clusclump"$k
C="clump1.clusclump"

s.sample input=site.clusclump output=$F rast=$F
s.to.rast input=$F output=$F size=0
echo $MESSAGE${LAYER[$k]} $F
k=`expr $k + 1`

done

MESSAGE="Clustering: "

k=2
com=", "
FF="caverage.clusclump1"
while [ $k -le $n ] ; do
F="caverage.clusclump"$k
FF=$FF$com$F
k=`expr $k + 1`

done

echo $MESSAGE$FF
r.clustering input=$FF output=$OUT k=$CL2 maxiter=$MAXIT

r.mapcalc $OUT="if(isnull($OUT),0,$OUT)"

r.fpstats dat=$OUT cove=clump1.clusclump stat=$OUT tota=ttt.clusclump

r.mapcalc $OUT="int($OUT*tt.clusclump)"

d.rast $OUT

#elimita files temporanei

k=2
com=", "
FF="caverage.clusclump1"
while [ $k -le $n ] ; do
F="caverage.clusclump"$k
FF=$FF$com$F
k=`expr $k + 1`

done

g.remove rast=$FF
g.remove rast=tt.clusclump,ttt.clusclump,clump1.clusclump,clust1.clusclump

g.remove sites=site.clusclump

```

## Versione che utilizza il fuzzy c-means

```
#!/bin/bash
```

```
declare -a LAYER ;  
declare -a CLAYER ;
```

```
nc=0
```

```
n=0
```

```
for i do
```

```
    case $i in
```

```
        clump1=*)    CLAYER[1]=`echo $i | sed s/clump1=//`  
                     nc=`expr $nc + 1` ;;  
        clump2=*)    CLAYER[2]=`echo $i | sed s/clump2=//`  
                     nc=`expr $nc + 1` ;;  
        clump3=*)    CLAYER[3]=`echo $i | sed s/clump3=//`  
                     nc=`expr $nc + 1` ;;  
        clump4=*)    CLAYER[4]=`echo $i | sed s/clump4=//`  
                     nc=`expr $nc + 1` ;;  
        clump5=*)    CLAYER[5]=`echo $i | sed s/clump5=//`  
                     nc=`expr $nc + 1` ;;  
        clump6=*)    CLAYER[6]=`echo $i | sed s/clump6=//`  
                     nc=`expr $nc + 1` ;;  
        clump7=*)    CLAYER[7]=`echo $i | sed s/clump7=//`  
                     nc=`expr $nc + 1` ;;  
        clump8=*)    CLAYER[8]=`echo $i | sed s/clump8=//`  
                     nc=`expr $nc + 1` ;;  
        clump9=*)    CLAYER[9]=`echo $i | sed s/clump9=//`  
                     nc=`expr $nc + 1` ;;  
        clump10=*)   CLAYER[10]=`echo $i | sed s/clump10=//`  
                     nc=`expr $nc + 1` ;;  
  
        layer1=*)    LAYER[1]=`echo $i | sed s/layer1=//`  
                     n=`expr $n + 1` ;;  
        layer2=*)    LAYER[2]=`echo $i | sed s/layer2=//`  
                     n=`expr $n + 1` ;;  
        layer3=*)    LAYER[3]=`echo $i | sed s/layer3=//`  
                     n=`expr $n + 1` ;;  
        layer4=*)    LAYER[4]=`echo $i | sed s/layer4=//`  
                     n=`expr $n + 1` ;;  
        layer5=*)    LAYER[5]=`echo $i | sed s/layer5=//`  
                     n=`expr $n + 1` ;;  
        layer6=*)    LAYER[6]=`echo $i | sed s/layer6=//`  
                     n=`expr $n + 1` ;;  
        layer7=*)    LAYER[7]=`echo $i | sed s/layer7=//`  
                     n=`expr $n + 1` ;;  
        layer8=*)    LAYER[8]=`echo $i | sed s/layer8=//`  
                     n=`expr $n + 1` ;;  
        layer9=*)    LAYER[9]=`echo $i | sed s/layer9=//`  
                     n=`expr $n + 1` ;;  
        layer10=*)   LAYER[10]=`echo $i | sed s/layer10=//`  
                     n=`expr $n + 1` ;;  
  
        class1=*)    CL1=`echo $i | sed s/class1=//` ;;  
        class2=*)    CL2=`echo $i | sed s/class2=//` ;;  
        nonoltre=*)  MAXIT=`echo $i | sed s/nonoltre=//` ;;  
  
        output=*)    OUT=`echo $i | sed s/output=//` ;;
```

```

        *)
            echo ""
            echo "Opzione non valida: " $i
        esac
    done

    k=2
    com=", "
    FF="{CLAYER[1]}"
    while [ $k -le $nc ] ; do
        FF=$FF$com${CLAYER[$k]}
        k=`expr $k + 1`
    done

    r.clustering input=$FF output=clust1.clusclump k=$CL1 maxiter=$MAXIT

    r.mapcalc clust1.clusclump="clust1.clusclump+1" > /dev/null

    r.clump -q input=clust1.clusclump output=clump1.clusclump


    k=1
    kk=1
    while [ $k -le $n ] ; do
        F="caverage.clusclump"$k
        C="clump1.clusclump"
        r.fpstats dat=${LAYER[$k]} cove=$C stat=$F tota=tt.clusclump
        echo $MESSAGE${LAYER[$k]}
        k=`expr $k + 1`
    done

    r.volume -q data=caverage.clusclump1 clump=clump1.clusclump
    site_list=site.clusclump >/dev/null

    MESSAGE="Converting: "
    k=1
    kk=1
    while [ $k -le $n ] ; do
        F="caverage.clusclump"$k
        C="clump1.clusclump"

        s.sample input=site.clusclump output=$F rast=$F
        s.to.rast input=$F output=$F size=0
        echo $MESSAGE${LAYER[$k]} $F
        k=`expr $k + 1`
    done

    MESSAGE="Clustering: "

    k=2
    com=", "
    FF="caverage.clusclump1"

```

```

while [ $k -le $n ] ; do
F="caverage.clusclump"$k
FF=$FF$com$F
k=`expr $k + 1`

done

echo $MESSAGE$FF
r.fuzzy input=$FF output=$OUT k=$CL2 maxiter=$MAXIT climt=0.0001

r.mapcalc $OUT="if(isnull($OUT),0,$OUT)"

r.fpstats dat=$OUT cove=clump1.clusclump stat=$OUT tota=ttt.clusclump

r.mapcalc $OUT="int($OUT*tt.clusclump)"

d.rast $OUT

#elimita files temporanei

k=2
com=", "
FF="caverage.clusclump1"
while [ $k -le $n ] ; do
F="caverage.clusclump"$k
FF=$FF$com$F
k=`expr $k + 1`

done

g.remove rast=$FF
g.remove rast=tt.clusclump,ttt.clusclump,clump1.clusclump,clust1.clusclump

g.remove sites=site.clusclump

```

## **Algoritmi implementati in linguaggio C come moduli di GRASS 5.0.1**

1. r.clustering: algoritmo a supporto dello script clusclump.sh che implementa il metodo K-means
2. r.fuzzy: algoritmo a supporto dello script clusclump.sh che implementa il metodo fuzzy c-means
3. r.fpstats: algoritmo a supporto dello script clusclump.sh. per estrarre dal raster dei “clump” i valori statistici
4. r.mvdist: algoritmo per generare il raster della distanza di Mahalanobis
5. r.terraces: algoritmo per la correlazione e l’ordinamento dei terrazzi fluviali (ed altre superfici relitte)